AD-A042 797 BELFER GRADUATE SCHOOL OF SCIENCE NEW YORK F/G 20/12
ELECTRON STATES NEAR SURFACES AND IN DISORDERED STRUCTURES.(U)
JUL 77 D C MATTIS N00014-76-C-0690

UNCLASSIFIED NL

1 OF 1
ADA042797

END
DATE
FILMED
9-77
DDC

OFFICE OF NAVAL RESEARCH

Contract No. N00014-76-C-0690

Task No. NR392-015

Report No. 2

ELECTRON STATES NEAR SURFACES AND IN DISORDERED STRUCTURES

by

Daniel C. Mattis

BELFER GRADUATE SCHOOL OF SCIENCE
Yeshiva University
New York, New York 10033

For Period: January 1 - July 31, 1977

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER DCM-B	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) Electron States near Surfaces and in Disordered Structures,		5. TYPE OF REPORT & PERIOD COVERED (9) Interim Report, no. 2, January 1 - July 31, 1977
		6. PERFORMING ORG. REPORT NUMBER 2 1 Jan - 31 Jul 77
7. AUTHOR(s) (10) Daniel C. Mattis		8. CONTRACT OR GRANT NUMBER(s) (15) N00014-76-C-0690
9. PERFORMING ORGANIZATION NAME AND ADDRESS Daniel C. Mattis, Belfer Graduate School of Science Yeshiva University, New York, New York 10033		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR392-015 (11) 31 Jul 77
11. CONTROLLING OFFICE NAME AND ADDRESS Department of the Navy, Office of Naval Research, 800 North Quincy Street, Arlington, Virginia 22217		12. REPORT DATE July 31, 1977
		13. NUMBER OF PAGES 72
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (12) 72p.		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approval for Public Release. Distribution Unlimited.

DDC
AUG 9 1977
C

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

1. Preprint of new work on theory of surfaces.
2. Reprints of paper on surfaces and disordered structures.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Electron states
Surface states
Random alloys

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The optical and transport properties of solids may be obtained from an accurate determination of the electron eigenstates. The three papers included in this report address themselves to the problem of obtaining such eigenstates under a variety of circumstances.

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

053820

1B

ELECTRON STATES NEAR SURFACES OF SOLIDS*

Daniel C. Mattis

BELFER GRADUATE SCHOOL OF SCIENCE
Yeshiva University
New York, New York 10033

ABSTRACT

The mathematical theory of electron eigenstates near the surfaces of solids is developed in stages. We first study the eigenstates of solids which terminate at a surface but are otherwise unperturbed, and prove that near the surface the 3-dimensional band edges are "softened" and van Hove's singularities in the density of states are eliminated. A set of "vacuum states" lying primarily outside the solid is constructed out of plane waves orthogonalized to the eigenstates of the solid. This set of vacuum states is not orthonormal, and must be orthonormalized by an original procedure, different from that of Schmidt (which is unsuited). Effects of surface perturbations are studied. An exact method is elaborated for obtaining eigenfunctions in closed form, for a variety of perturbing potentials, extending arbitrary distances from the surface and including inter-band matrix elements. It consists of calculating the effects of one surface layer at a time, and cumulating the results. The intrinsic instability of certain surfaces against the formation of bands of surface states is shown to be the consequences of the vanishing of a bulk quantity α_{zz}, one of the components of the inverse-effective mass tensor, at certain points in the B.Z.

*Research supported in part by the Office of Naval Research, and the National Science Foundation.

I. INTRODUCTION

With a vast literature now appeared on various aspects of the theory of surfaces of solids, it might seem that the scope for further investigation is thoroughly circumscribed. Yet, even a cursory glance at this published literature reveals the lack of strong mathematical underpinnings, of the type that Floquet's theory and group theory gave to the electron states (the Bloch functions) of the periodic solid. Many ~~Much~~ of the published works consist principally of detailed numerical investigations of the Schrödinger differential equation in the neighborhood of the surfaces. There, perturbing potentials, both one-body and two-body, are taken into account, self-consistently in the best of these investigations. The study of easily solved simple models, such as nearest-neighbor simple cubic ("cubium") structures, yields only qualitative "seat of the pants" experience but no generally applicable formulas or theorems.

In the present work we have taken an approach which we hope the reader will agree to be justified and rewarding. The ~~Our~~ idea is that in any surface problem, the proper basis set in which to do the calculations must be adapted to the symmetry of the problem, in this case the termination of the solid. If this feature is "built in", ~~that~~ many features common to all surfaces can be identified in advance, and the calculations for specific problems reduced to a minimum.

Then, in studying the effects of surface potentials on the~~se~~ surface-adapted functions, we find these ~~that they are~~ now quite muted and predictable; when the perturbing potential is sufficiently strong or the energy contours sufficiently flat (the $\alpha_{zz} = 0$ criterion[1] — see below),

there appear bound states, that are identified as surface states. In general, the basis functions are scattered among themselves by the perturbations at the surface. Calculations of the scattering and of the bound state properties can be reduced to straightforward simple algebra, and the introduction of more sophisticated effects (many-body correlations, principally) then involve no more difficult concepts or calculations than for the bulk. While not a universal theory of surface of solids, the present work does outline procedures that re-cycle to a maximum extent the considerable information, already available about periodic solids, to the study of the surface related properties. Let us now examine some of the concepts that are involved. ~~help bring this about~~.

Although chemists describe molecules as a collection of atoms interconnected by "chemical bonds", these being quantities which retain their integrity (length, strength, and relative angles) in a variety of compounds, little use of this important concept has been made in solid-state theory outside the framework of such approximate schemes such as LCAO. In the present paper, we make the obvious identification of the bond $K(R_{ij})$ connecting an atom at R_i to a second atom at R_j as the lattice Fourier transform of the band energy $\epsilon_n(k)$. The exact Bloch energies $\epsilon_n(k)$ and corresponding eigenstates $\psi_{nk}(r)$ of the periodic lattice are the principal ingredients in the present formulation of the electronic theory of surfaces of solids.[1]

(A) Our theory of surfaces is developed in several stages. As the first and most important step, bonds which connect the "inside" atoms to "outside" atoms are cut at a surface introduced into an otherwise

semi-infinite solid; internal bonds as well as the internal Hamiltonian retain their bulk values. The exact new eigenstates are then simply obtained as linear combinations of the Bloch functions, the coefficients being determined by a set of boundary conditions. While this procedure can be carried out explicitly ~~analytically~~ only for extremely simple examples of nearest neighbor or next-nearest neighbor bonding,[2] or for an empty lattice, our method is so formulated that in general, exact eigenstates at this stage can be easily computed ~~numerically~~, using simple matrix algebra on the known bulk solutions. Moreover, we derive a variety of properties of the eigenstates of the terminated solid without recourse to numerical methods. For example, we prove in all generality:

-There are no "surface states", i.e. no bound states lying in the ranges of energy forbidden to bulk states. The energy eigenvalues of the eigenfunctions of the terminated but otherwise unperturbed solid interlace the Bloch energies, from which they differ only to order $O(1/N_z)$, where N_z = number of planes of atoms parallel to the surface $\to \infty$.

-At energies $\epsilon_n(k)$ such that the velocity of an electron perpendicular to the surface $v_z = \hbar^{-1}\, \partial\epsilon_n(k)/\partial k_z$ vanishes, so does the relative probability $|\psi_{nk}(k)|^2$ of finding it at a finite distance z from the surface. Thus the local density of states (l.d.o.s.) at planes near the surface are lacking the accumulated contributions from band edges and saddle points in the energy spectrum, with consequent erosion of the band edges and elimination of the van Hove singularities characteristic of the bulk d.o.s. and band structure.[3]

(B) Turning to the second stage of the theory, the instabilities of the new eigenfunctions against simple model surface perturbations is examined. Generally, surface potentials, or changes in the bond strengths, can "pull" eigenstates into the surface or "push" them out of it. More importantly, such perturbations can cause bound states, i.e. surface states, to appear ~~outside~~ alongside the continuum of the Bloch states. Thus, for a given point k_x, k_y in the two-dimensional (2D) Brillouin Zone (B.Z.) of the surface, we could have states lying outside the continuum of energies $\epsilon_n(k_x,k_y;k_z)$ (defined by allowing $k_z a$ to wander over the interval $-\pi$ to $+\pi$), provided the perturbation is sufficiently strong or the incipient instability sufficiently great. At this point we show that the instabilities arise <u>at or near certain unusual points in the 2D B.Z.</u> where the inverse-effective-mass parameter $\alpha_{zz}(k_x,k_y) \equiv \hbar^{-2}\, \partial^2 \epsilon_n(k)/\partial k_z^2$ vanishes, the derivative being evaluated at the values of k_z for which $\epsilon_n(k_x,k_y;k_z)$ has its maximal and/or minimal values, at fixed n, k_x and k_y. Only a knowledge of the bulk band structure and a choice of surface orientation relative to the crystal axes (to determine the orientation of the z-axis, normal to the surface) are required to locate the zeros of α_{zz}, and thus to establish the bands and surface orientations most likely to produce surface states.[1]

(C) In the third stage, more general perturbations are considered and the mathematical procedure for obtaining the new eigenstates and eigenvalues — both in the continuum and for the bound surface states — is developed. It is only for large perturbations that we can expect to see surface states lying within the bulk band-gaps, as weak perturbations produce surface states whose energies usually overlap parts

of the bulk continuum, but do not mix as long as they belong to different values of n or k_x and k_y. The various contributions to the perturbing Hamiltonian include: a changed Madelung potential (zero outside the solid, varying progressively to its bulk limiting value at large z), a macroscopic dipole moment due to the imbalance of forward-backward charge on orbitals of surface atoms, relating to changes to the usual Hartree-Fock and correlation energies due to termination of the crystal.[4] In addition to scattering of the various eigenstates, such perturbations can also mix different bands (conserving k_x, k_y however), resulting in a distortion (rotation, elongation, or whatever) of the original bonds. They can also admix the "vacuum states", originally decoupled by the breaking of surface bonds.

For the purposes of a realistic calculation, one must therefore reintroduce a complete set of states for the outside space — the "vacuum levels". These may be chosen as a complete set of plane wave states displaced by an energy determined by the magnitude of the work function, and orthogonalized to the eigenstates of the solid near the surface. A procedure for doing this is outlined. The perturbations described in the previous paragraph have matrix elements connecting the inside states to the vacuum states as well as to each other, and so we had to have a scheme for calculating them.[5]

This work is intended to provide a thorough introduction to a variety of new mathematical aspects only; we intend to present applications in subsequent publications. Problems concerning "reconstruction" of surfaces are also deferred.

II. SOME FAMILIAR BUT USEFUL CONCEPTS

The bulk Schrödinger equation $H\psi(r) = \epsilon\psi(r)$ determines the complete, orthonormal set of Bloch states ψ . These can be written in two distinct ways:

$$\psi_{nk}(\vec{r}) = N_{x,y,z}^{-\frac{1}{2}} \left[e^{i\vec{k}\cdot\vec{r}} \underline{u}_{nk}(\vec{r})\right] \text{ or}$$

$$\psi_{nk}(\vec{r}) = N_{x,y,z}^{-\frac{1}{2}} \sum_{j} e^{i(\vec{k}\cdot\vec{R}_j)} \phi_n(\vec{r}-\vec{R}_j), \tag{2.1}$$

in which $u_{nk}(\vec{r})$ is a function having periodicity of the lattice, and $\phi_n(\vec{r}-\vec{R}_j)$ is the Wannier function for the nth band at the R_jth lattice site. The orthonormal Wannier function is localized within a small neighborhood of its nominal site, decaying exponentially at large distances from it. The energy eigenvalue of the above Bloch state is:

$$\epsilon_n(k) = \frac{1}{N_{x,y,z}} \sum_{\vec{R}_i} \sum_{\vec{R}_j} e^{ik\cdot\vec{R}_{ij}} K_n(\vec{R}_{ij}) \tag{2.2}$$

It is, in the extended zone scheme, periodically extensible into the second, third, etc. Brillouin Zones. For our purposes, knowledge of (2.2) for $\vec{k}$ in the first Brillouin Zone is sufficient. For the study of surfaces, an even smaller, two-dimensional, Brillouin Zone is important. Let the normal to the surface plane define the z-direction, so that the surface defines the x-y plane. The new boundary conditions at the surface and the perturbing potentials localized near the surface may or may not affect the translational symmetry in the x-y plane. When they do, we say the surface is "reconstructed".[6] In the present work we avoid these complications. We shall always assume the surface

to be unreconstructed, i.e. we shall make strong use of the hypothesis that the translational symmetry in the x-y plane is maintained exactly as in the bulk even though this symmetry is destroyed in the z-direction by the surface.

At every fixed value of $k_{\parallel} \equiv (k_x, k_y)$ the energy $\epsilon(k_{\parallel}, k_z)$ achieves a minimum and a maximum, at $k_{z\,min}$ and $k_{z\,max}$ respectively. These will be denoted the "relative band edges" to distinguish them from the "absolute band edges" at which the entire nth band achieves its lowest and highest energies. We denote the lower relative band edge $\epsilon_n^{min}(k_{\parallel})$ and the upper one $\epsilon_n^{max}(k_{\parallel})$.

At any k in the nth band one defines an inverse-effective-mass tensor $\overset{\rightleftarrows}{\alpha}_n$ with nine components $\alpha_{n,ij}$ given by:

$$\alpha_{n,ij}(\vec{k}) = (\hbar)^{-2} \frac{\partial^2}{\partial k_i \partial k_j} \epsilon_n(\vec{k}) \tag{2.3}$$

In surface studies, one is particularly concerned with just one of these components, $\alpha_{n,zz}$, evaluated at $k_{z\,min}$ and $k_{z\,max}$. We introduce the self-explanatory notation $\alpha_{n,zz}^{min}(k_{\parallel})$ and $\alpha_{n,zz}^{max}(k_{\parallel})$, both functions only of $k_{\parallel}$; quantities which, together with $\epsilon_n^{min}(k_{\parallel})$ and $\epsilon_n^{max}(k_{\parallel})$ we must evaluate using knowledge of the usual (bulk-)band structure.

To fix ideas, consider the hypothetical simple-cubic tight-binding "cubium", which has a band structure:

$$\epsilon_n = K_n(\cos k_x a_o + \cos k_y a_o + \cos k_y a_o) \tag{2.4a}$$

for the surface perpendicular to one of the crystal axes, and

$$\epsilon_n = K_n(\cos k_x a_o + 2\cos k_y a_o 2^{-\frac{1}{2}} \cos k_z a_o 2^{-\frac{1}{2}}) \qquad (2.4b)$$

for a surface in the (0,1,1) direction. For the first, (0,0,1) orientation, we obtain assuming $K < 0$,

$$k_{z\ min} = 0,\ k_{z\ max} = \pi/a,\ \epsilon^{min,max} = K_n(\cos k_x a_o + \cos k_y a_o) \pm K_n$$

$$\text{and } \alpha_{zz}^{min,max} = \pm |K_n| a_o^2 \text{ (setting } \hbar = 1),$$

whereas for the second, (0,1,1) orientation,

$$k_{z\ min} = 0 \text{ for } \cos(k_y a_o 2^{-\frac{1}{2}}) > 0 \text{ and } k_{z\ min} = \pi 2^{\frac{1}{2}}/a_o \text{ for } \cos(k_y a_o 2^{-\frac{1}{2}}) < 0,$$

with $k_{z\ max} = \pi 2^{\frac{1}{2}}/a_o$ in the former and 0 in the latter case. Then,

$$\epsilon^{min,max} = K_n(\cos k_x a_o \pm 2\,|\cos k_y a_o 2^{-\frac{1}{2}}|)$$

and

$$\alpha_{zz}^{min,max} = \pm |K_n a^2 \cos k_y a_o 2^{-\frac{1}{2}}| \qquad (2.5)$$

For the first orientation, the two-dimensional Brillouin Zone (2DBZ) is a square extending from $-\pi$ to $+\pi$ for both $k_x a_o$ and $k_y a_o$. For the second, it is a rectangle, with $k_x a_o$ extending from $-\pi$ to $+\pi$ while $k_y a_o$ extends from $-\pi 2^{\frac{1}{2}}$ to $+\pi 2^{\frac{1}{2}}$. We display contours of constant ϵ^{min} and the behavior of α_{zz}^{min} in Figures 1 and 2 (ϵ_{max} and α^{max} are similar).

In equations (2.1) and (2.2) the factors $N_{x,y,z}$ refer to the total number of cells in the 3D crystal. With N_z the total number of planes

parallel to the surface plane, and $N_{x,y}$ the number of cells in each plane, we have obviously:

$$N_{x,y,z} = N_z N_{x,y} \tag{2.6}$$

$K_n(\vec{R}_{ij})$ in equation (2.2) is the "bond strength" or Hamiltonian matrix-element connecting the Wannier function in the nth band in the cell centered at $\vec{R}_i$ to that at $\vec{R}_j$. While in the periodic solid K_n is a function only of the relative vector $\vec{R}_{ij} \equiv \vec{R}_i - \vec{R}_j$, introduction of a surface complicates matters. Zero Boundary Conditions require that we cut all bonds that extend outside the solid, i.e. require us to set all $K_n = 0$ for which either $\vec{R}_i$ or $\vec{R}_j$ lie outside the solid. This procedure will be examined in the following Section.

In the tight-binding scheme[7] the bond-strengths $K_n(\vec{R}_{ij})$ are estimated first, and the band-structure calculated therefrom. In more modern schemes,[8] the band structure $\epsilon_n(\vec{k})$ is obtained by direct solution of Schrödinger's differential equation appropriate to the periodic solid, with or without correlation and exchange corrections. Once this is done, we can convert (2.2) to obtain the K_n's:

$$K_n(\vec{R}_{ij}) = \left(\frac{a_o}{2\pi}\right)^3 \int_{1^{st}B.Z.} d_3k \; e^{-i\vec{k}\cdot\vec{R}_{ij}} \epsilon_n(\vec{k}) \; . \tag{2.7}$$

In the Wannier-function representation, the eigenstates of the Hamiltonian for electrons in the solid take the general form:

$$\psi(r) = \sum_j U(\vec{R}_j) \phi_n(\vec{r}-\vec{R}_j) \tag{2.8}$$

where the U's are the components of the eigenvectors of the N x N matrix (2.7), and are to be normalized to unit length. By using a translation-operator representation of (2.7) we can obtain a Schrödinger-like equation for these amplitudes:

$$H_n\, U(\vec{R}) = E\, U(\vec{R}) \tag{2.9}$$

in which the "Hamiltonian" is a generalized translation operator:

$$H_n \equiv \epsilon_n(-i\partial/\partial\vec{R}) = \sum_{\vec{R}_j} K_n(\vec{R}_j)\, \exp(R_j \cdot \partial/\partial\vec{R}) \tag{2.10}$$

When subject to PBC the normalized solutions of (2.9) are evidently plane waves $N_{x,y,z}^{-\frac{1}{2}}\, \exp(i\vec{k}\cdot\vec{R}_j)$ and (2.8) reduces to the Bloch form (2.1), with the energy eigenvalue $E = \epsilon_n(\vec{k})$.

In tight-binding, the Wannier functions are presumed known, or are approximated by the corresponding atomic orbitals.[7] In the more modern procedures, the Bloch states or their periodic components $\underline{u}_{nk}(\vec{r})$ are obtained by a direct computer solution of the appropriate Schrödinger differential equation;[8] we may then use these known eigenfunctions to construct the Wannier functions, by inverting (2.1):

$$\emptyset_n(\vec{r}-\vec{R}_j) = N_{x,y,z}^{-\frac{1}{2}} \sum_{k \subseteq 1^{st}B.Z.} e^{-i\vec{k}\cdot\vec{R}_j}\, \psi_{nk}(\vec{r})$$

$$= \left(\frac{a_o}{2\pi}\right)^3 \int_{1^{st}B.Z.} d_3k\; e^{ik\cdot(r-R_j)}\, \underline{u}_{nk}(\vec{r}) \tag{2.11}$$

We shall henceforth assume these functions to be known.

III. ZERO BOUNDARY CONDITIONS

A terminated solid consists of atomic planes at $Z_j = m_j a$, with $m_j = 1,2,\ldots,N_z$. The interplanar distance a is a function of the orientation of the crystal axes, of course ($a = a_o$ for the (001) example in Section II and $a = a_o 2^{-\frac{1}{2}}$ in the second, (011), example). The existence of a surface at $m_j = 0$ eliminates the PBC in favor of requirements that the amplitudes U (defined in the preceeding Section) vanish at $m_j = 0,-1,-2,\ldots$. These are denoted the zero boundary conditions ZBC.

We take advantage of periodicity in the x-y plane to factor the amplitudes:

$$U_{nk}(\vec{R}_j) = N_{x,y}^{-\frac{1}{2}} \exp i(k_x X_j + k_y Y_j)\, U_{nk}(Z_j) \tag{3.1}$$

where the U(Z)'s satisfy the difference equation:

$$\mathcal{H}\, U_{nk}(Z) = E\, U_{nk}(Z) \tag{3.2}$$

for which the translation operator is given as:

$$\mathcal{H} \equiv \epsilon_n(k_{\|}, \tfrac{1}{i}\partial/\partial Z) \tag{3.3}$$

Alternately,

$$\mathcal{H} = \sum_{p=-P}^{+P} K(p)\, e^{pa\partial/\partial Z} \tag{3.4}$$

in which

$$K(p) = \frac{a}{2\pi} \int_{-\pi/a}^{+\pi/a} dk_z \; \epsilon_n(k_{\|},k_z) \; e^{-ipak_z}, \tag{3.5}$$

by obvious specializations of the preceeding Section. It is important to emphasize that the coefficients $K(p)$ will vary with planar wave-vector $k_{\|} \equiv (k_x,k_y)$ and band index $\underline{n}$ and should be so labelled: $K_{n,k_{\|}}(p)$. For all reasonable band structures, one may without significant error cut off the sum in (3.4) beyond a finite maximum integer P, which rarely exceeds 0(10) but may often be as small as 1 or 2. We assume the existence of such a cut-off P, and amplify much of the material in reference 1 for the sake of a complete presentation.

We start with the inverse of (3.5):

$$\epsilon(k_z) = \sum_{p=-P}^{+P} K(p)(e^{ik_z a})^p \tag{3.6}$$

Here and elsewhere in this paper, we suppress both the band index $\underline{n}$ and the planar wave-vector $k_{\|}$ for reasons of typographical clarity. Wherever this could occasion some confusion, we restore all the parameters. We also use the notation $\tau \equiv (k_{\|},n)$ on occasion, e.g. $K_\tau(p)$ is the quantity defined in (3.5) and $\epsilon_\tau(k_z)$ in (3.6), the latter also denoted $\epsilon_n(\vec{k})$.

The equation (3.6) will now be viewed as a 2Pth order equation for $\zeta \equiv \exp(ik_z a)$, having in general, 2P distinct complex roots. We label these

$\zeta_\alpha = \exp(ik_z^{(\alpha)}a)$, with index $\alpha = 1,2,\ldots,2P$. The most general solution of (3.2) at a fixed $E = \epsilon(k_z)$ for a real specified value of k_z is thus:

$$U(Z_j) = \sum_{\alpha=1}^{2P} A_\alpha \exp(ik_z^{(\alpha)}m_j a) = \sum_{\alpha=1}^{2P} A_\alpha \zeta_\alpha^{m_j} \tag{3.7}$$

in which the given k_z is one of the members of the set of $k_z^{(\alpha)}$'s, e.g. $k_z = k_z^{(1)}$. The coefficients A_α are now obtained, under the simplifying assumptions of no-spin-orbit coupling, which allows us to assume the K's in (3.4) - (3.6) real and symmetric under the interchange $p \leftrightarrow -p$.

(i) By an obvious symmetry, for every root k there is a second root -k. We know one real root, $k_z^{(1)}$ = the specified k_z. Including its negative, there is a total of 2Q real roots, with $Q \geq 1$. The remaining 2(P-Q) roots are complex, so half of them (those with negative imaginary parts) must be discarded because they lead to growing (and therefore un-normalizable) exponentials inside the solid ($m_j > 0$). The magnitudes are: $|\zeta| = 1$ for 2Q real-k_z roots and $|\zeta| < 1$ for the remaining P-Q complex-k_z roots that are retained.

(ii) An alternate formulation of (3.6) regards this as a Pth order polynomial equation in $\cos(k_z a)$. Because of the real coefficients, the roots come in complex pairs such as:

$$(\cos k_z a - u + iw)(\cos k_z a - u - iw) = 0 \tag{3.8}$$

If $k_z^{(\alpha)}$ is here an acceptable root (i.e. Im $k_z^{(\alpha)} \geq 0$), then, by inspection, so is $-k_z^{(\alpha)*}$, and the two are distinct unless they are purely imaginary. Summarizing: the

(a) purely real $k_z^{(\alpha)}$'s come in $\pm$ pairs, the

(b) complex $k_z^{(\alpha)}$'s come in pairs: $k_z^{(\alpha)}$ and $-k_z^{(\alpha)*}$, while the

(c) Pure imaginary $k_z^{(\alpha)}$'s are necessarily unique.

(iii) We now impose ZBC on (3.7), setting $U(Z_j) = 0$ for $m_j = 0,-1,\ldots,-(P-1)$. There is no requirement at $m_j \leq -P$ as the bonds within the solid, which are to be "broken" by the imposition of the boundary conditions, extend out only as far as $-(P-1)$. Thus the $P + Q$ non-zero coefficients A_α in the expansion (3.7) are constrained by a set of P linear homogeneous equations, and are reduced to a set of at most Q independent members. The net number Q is determinable as follows: we study the behavior of $\epsilon(k_z)$ as k_z is varied in the real interval $0 \leq k_z \leq \pi/a$. In the range where ϵ is single-valued, $Q = 1$. In the range (if any) where ϵ is double-valued, $Q = 2$, etc. This is illustrated in Figure 3. We note that the net number of solutions is precisely half of that obtained using PBC, a not unexpected result.

(iv) Orthogonalization and Normalization of solutions. We distinguish two types of sums, those which depend on N_z:

$$\sum_{m=1}^{N_z} |\zeta_\alpha|^{2m} = N_z, \ \underline{\text{only for}}\ k_z^{(\alpha)}\ \underline{\text{real}},\ \text{i.e. } |\zeta_\alpha| = 1 \tag{3.9}$$

and those which, in the thermodynamic limit $N_z \to \infty$, are independent of N_z:

$$\sum_{m=1}^{N_z} (\zeta_\alpha^* \zeta_\beta)^m = \sum_{m=1}^{\infty} (\zeta_\alpha^* \zeta_\beta)^m = \frac{\zeta_\alpha^* \zeta_\beta}{1-\zeta_\alpha^* \zeta_\beta} \tag{3.10}$$

The latter occur when either $k_z^{(\alpha)}$ or $k_z^{(\beta)}$ are complex or imaginary, or, in general, for $\alpha \neq \beta$. (We recall that the imaginary parts are always positive; for <u>real</u> $k_z^{(\alpha)}$ we may here (but not in (3.9)) introduce an infinitesimal positive imaginary part to speed the convergence.)

(a) Let us now normalize a function of the type (3.7) when $Q = 1$. There are $P + 1$ terms in the expansion (3.7): the specified k_z and $-k_z$, a certain number S of purely imaginary $k_z^{(\alpha)}$'s, and the remainder (P-1-S) complex pairs of the type $k_z^{(\alpha)}$ and $-k_z^{(\alpha)*}$. Thus, (P-1-S) is an even integer or zero. We write U as follows:

$$U(Z_j) = A_1(\zeta_1^{m_j} + a_2\zeta_2^{m_j} + a_3\zeta_3^{m_j} + \ldots) \tag{3.11}$$

The ratios $a_2 = A_2/A_1$, $a_3 = A_3/A_1$,... are determined by the P boundary conditions. We have chosen $k_z^{(1)} = k_z$, $k_z^{(2)} = -k_z$, so that $|\zeta_1| = |\zeta_2| = 1$ and all other $|\zeta_i| < 1$. In part (b) below, we shall prove $A_1 \neq 0$.

Because both the Hamiltonian and the boundary conditions are real the solutions (3.7) are real (or can be made real by a trivial (constant) multiplicative factor). Thus, $a_2A_1 = A_1^*$; the coefficients a_iA_1 for the S terms involving the pure imaginary $k_z^{(\alpha)}$'s are all real; and the coefficients a_n, a_{n+1} of the (P-1-S) conjugate pairs are themselves complex conjugates: $a_nA_1 = a_{n+1}^*A_1^*$.

The normalization requirement is thus

$$1 = \sum_{m_j=1}^{N_z} |A_1|^2 |[\zeta_1^{m_j} + a_2 \zeta_2^{m_j} + \ldots]|^2$$

$$= |A_1|^2 [2N_z + O(1)] \tag{3.12}$$

by use of (3.9) and (3.10) and $|a_2|^2 = 1$. Thus, A_1 is a constant:

$$|A_1| = (1/2N_z)^{\frac{1}{2}} \tag{3.13}$$

independent of $\varepsilon(k_z)$ and of a_i for $i \geq 3$, a fortunate simplification.

(b) The degenerate cases for $Q \geq 2$ are treated analogously, by constructing Q linearly independent functions of the type (3.11) each containing one or more of the real k_z's (and their mates $-k_z$), orthogonalizing them, and then normalizing each.

This requires us to prove that one cannot construct a solution utilizing only the P complex or imaginary allowed k_z's. Such a proof is also required in part (a) preceeding, to justify the assumption $A_1 \neq 0$, on which the entire analysis (3.11) - (3.13) was based.

Supposing one could construct a solution U using any subset ($P' \leq P$) of P' distinct $k_z^{(\alpha)}$'s, labeled $\alpha = Q+1, Q+2, \ldots, Q+P'$, such as:

$$U(m) = \sum_{r=1}^{P'} A_r \zeta_{Q+r}^m \tag{3.14}$$

with the ZBC applied at $m = 0, -1, \ldots, -(P'-1)$ to determine the A_r's. These P' equations may be conveniently summarized in matrix form:

$$\begin{array}{cc} & r \rightarrow \\ s\downarrow & \begin{bmatrix} & & \\ & \tilde{\tilde{M}} & \\ & & \end{bmatrix} \end{array} \cdot \begin{bmatrix} A_1 \\ A_2 \\ \cdot \\ \cdot \\ \cdot \\ A_{P'} \end{bmatrix} = 0 \qquad (3.15)$$

in which the elements of the $P'xP'$ matrix are:

$$M_{rs} = \zeta_r^{1-s}, \text{ column index, } r = 1,\ldots,P'$$
$$\text{and row index } s = 1,\ldots,P' \qquad (3.16)$$

Equation (3.15) can be satisfied iff the determinant of $\tilde{\tilde{M}}$ vanishes. This is recognized as Vandermonde's determinant, having the value:

$$\text{Det } \|\tilde{\tilde{M}}\| = \prod_{q>r}^{P'} \prod_{r=1}^{P'-1} (\zeta_q^{-1} - \zeta_r^{-1}) \qquad (3.17)$$

which cannot vanish as the ζ_r's are all distinct, Q.E.D.

The above has as one important consequence, that no decaying solution (alias surface-state) can be found to the unperturbed ZBC problem, i.e. all states in this model have at least one component with $|\zeta| = 1$ that survives in the bulk. A second corollary is that no new states are created in the energy-gap region by the imposition of ZBC, as $Q = 0$ there.

The next step is to construct Q distinct functions of type (3.11), each leading off with a different ζ_i chosen from the set of Q real, positive, specified k_z's illustrated in Figure 3. The reals again come in pairs: thus $|a_2| = 1$; however we can no longer assume that $k_z^{(3)}$ and

the succeeding ones are complex, as a number of them may belong to the set of reals. Let $R_1\{\alpha\}$ denote the set of reals contained in one of the linearly independent solutions; it contains at least 2 members, and at most 2Q. Let R_2 be the set of reals in a second solution. The sets R_1 and R_2 may be distinct, in which case, by equations (3.9) and (3.10) the two solutions are effectively orthogonal (in the limit $N_z \to \infty$). However, they may have one or more pairs of reals in common, as in the following example:

$$U_1 = A(\zeta_1^m - \zeta_2^m + a_3\,\zeta_3^m + a_3^*\,\zeta_4^m + \ldots)$$

and

$$U_2 = B(\zeta_3^m \pm \zeta_4^m + \ldots),$$

the dots indicating remaining terms all involving complex or imaginary k_z's. The normalization requirement for U_2 leads to:

$$|B|^2 = [2N_z]^{-1}.$$

We Schmidt-orthogonalize the two solutions, replacing U_1 by:

$$U_1' = U_1 - (N_z A^* B(a_3^* \pm a_3))U_2$$

The solution U_1' can now also be normalized. Note that the Schmidt-orthogonalization and the normalizations are all to be carried out without regard to evanescent complex or imaginary k_z's. This may be regarded as one of the important simplifications afforded by proceeding to the infinite N_z "thermodynamic" limit, and distinguishes the

present procedure from those[9] in which the solid is modeled by a finite number of parallel planes.

(v) Deep in the bulk ($m_j \to \infty$) the evanescent components disappear, so for any properly normalized (and, if need be, Schmidt-orthogonalized) solution $U(m)$:

$$\lim_{m \to \infty} |U(m)|^2 = 1/N_z \tag{3.18}$$

at any energy $\epsilon(k_z)$, for any degree Q of degeneracy. We define the various local density-of-states functions[10] for the mth plane in the nth band as:

$$\rho^o_{n,m}(E) \equiv N^{-1}_{xy} \sum_{k_{||}} \rho^o_{\tau,m}(E), \text{ where } \tau = (n,k_{||}) \text{ and}$$

$$\rho_{\tau,m}(E) \equiv \sum_{k_z} |U_{nk}(m)|^2 \, \delta(E-\epsilon_n(k_{||},k_z)) \tag{3.19}$$

By the preceeding, we see that the bulk limit $\rho^o_\tau(E)$ is attained:

$$\lim_{m \to \infty} \rho^o_{\tau,m}(E) \equiv \rho^o_\tau(E) = N_z^{-1} \sum_{k_z} \delta(E-\epsilon_n(k_{||},k_z)) \tag{3.20}$$

which includes all degeneracies in the nth band. The summed density-of-states $\rho(E) \equiv \Sigma\rho_n(E)$ is sometimes also of interest. The evaluation of such quantities at or near the surface, where the evanescent contributions are in principle important, is the topic of Section 5.

IV. SURFACE AMPLITUDES (UNPERTURBED)

The nearest-neighbor "tightest-binding" (P=1) model[2,7,11] is readily solved by the above procedures,[11] to obtain:

$$U_k(m) = (2/N_z)^{\frac{1}{2}} \sin(k_z\ am) \tag{4.1}$$

The bulk limits (3.18) - (3.20) can be explicitly verified, and the behavior near the surface can be determined. In particular, we see that at finite distance $\underline{m}$ from the surface the amplitudes near the band edges (at $k_z = 0$ and π/a) vanish as $(k_z - k_{z\ min})$ and $(k_z - k_{z\ max})$ respectively.

In this Section we shall prove, confirming our earlier conjecture,[1] that this "erosion" of the band edges is a perfectly general phenomenon. We shall then analyze the behavior of the surface density-of-states function and of such important quantities as $U_k^*(m)U_{k'}(m')$ near the relative band edges, as required for the analysis of the perturbed surfaces in the following Sections.

For the sake of definiteness, we study only the band minima; the maxima are studied in precisely the same way, and do not require additional commentary. Note in passing that a relative band extremum can be non-degenerate (Q=1) only if it occurs at $k_z = 0$ or π/a as in (4.1) above. If it occurs at an arbitrary point within the Zone, $Q \geq 2$; a local minimum such as point (k_b) in Figure 3 has $Q \geq 3$.

(a) To start, let Q = 1, P arbitrary. We use equations (3.11) - (3.13) for the form of $U_k(m)$. We now write the P "ZBC equations" for m = 0,-1,..., plus one equation at m = 1 for $U_k(1)$, extending the matrix formulation introduced in equation (3.15):

$$\left[\quad \underset{\approx}{M}' \quad\right] \cdot \begin{bmatrix} A_1 \\ A_2 \\ \cdot \\ \cdot \\ \cdot \\ A_{P+1} \end{bmatrix} = \begin{bmatrix} U_k(1) \\ 0 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \tag{4.2}$$

in which the P+1 dimensional square matrix M' has elements:

$$M'_{rs} = \zeta_r^{2-s} \tag{4.3}$$

We know from prior analysis $|A_1| = (1/2N_z)^{\frac{1}{2}} = |A_2|$, (see equations (3.11) - (3.13)). We now wish to obtain rigorous bounds on $U_k(1)$ and $A_3, A_4, \ldots, A_{P+1}$, in the limit as k_z approaches $k_{z\,min}$. Cramer's rule is used to invert (4.2):

$$A_r = D_r / \mathrm{Det}\|M'\| \tag{4.4}$$

in which D_r is the determinant of a matrix, of which the column vector of the r.h.s. of (4.2) is the rth column, while all other columns are the same as in M'. Explicitly,

$$A_1 = \frac{\begin{Vmatrix} U_k(1) & \zeta_2 & \zeta_3 \cdots \\ 0 & 1 & 1 \cdots \\ 0 & \zeta_2^{-1} & \zeta_3^{-1} \cdots \\ \vdots & & \end{Vmatrix}}{\begin{Vmatrix} \zeta_1 & \zeta_2 & \zeta_3 \cdots \\ 1 & 1 & 1 \cdots \\ \zeta_2^{-1} & \zeta_2^{-1} & \zeta_3^{-1} \cdots \\ \zeta_1^{-2} & \zeta_2^{-2} & \zeta_3^{-2} \\ \vdots & \vdots & \vdots \end{Vmatrix}} \tag{4.5}$$

The numerator factors into $U_k(1)$ times a Vandermonde determinant of the type (3.17), denoted F_1. In the limit $k_z \to k_{z\,min}$, when $\zeta_1 \to \zeta_2$, the denominator (also a Vandermonde determinant) vanishes as a linear power:

$$\lim_{k_z \to k_{z\,min}} \mathrm{Det}\|M'\| = (\sin k_z a) C_1 \tag{4.6}$$

as the first two columns become identical and $\zeta_2 = \zeta_1^{-1} = \exp{-ik_z a}$, with

$k_{z\,min} = 0$ or π/a. C_1 is a constant near $k_{z\,min}$. Thus, substituting the known value of A_1 yields:

$$\lim_{k_z \to k_{z\,min}} |U_k(1)| = |\sin k_z a| \; |C_1/F_1| (1/2N_z)^{\frac{1}{2}} \longrightarrow 0 \qquad (4.7)$$

The vanishing of $U_k(1)$ in the limit is ensured by the non-vanishing of the Vandermonde determinant F_1.

In the limit $k_z \to k_{z\,min}$, the analogous expression for A_2 differs from (4.5) only by interchanging of the second and first columns in the numerator. This establishes a second result:

$$\lim_{k_z \to k_{z\,min}} A_2 = -A_1 = i(1/2N_z)^{\frac{1}{2}}, \qquad (4.8)$$

the factor of $\underline{i}$ is to make U real. Finally, for $r \geq 3$ we have $D_r = U_k(1) \times F_r$, with F_r appropriate Vandermonde determinant, and so:

$$\lim . |A_r| = (1/2N_z)^{\frac{1}{2}} |F_r/F_1| = (1/2N_z)^{\frac{1}{2}} B_r \sin k_z a \to 0 \qquad (4.9)$$

in which B_r is constant in the neighborhood of $k_{z\,min}$. We have made use of the fact that for $r \geq 3$, $F_r \propto (\zeta_1 - \zeta_2) = 2i \sin k_z a$, with $k_z \to 0$ or π/a.

This last result (4.9) is important in that it sharply limits the extent to which the evanescent (decaying) waves contribute to the solution at the extremum. It establishes the qualitative similarity between all the Q=1 type extremum solutions, for arbitrary band structure, with one another and with the simplistic P=1 limit, equation (4.1).

(b) Turning to the more complicated analysis of $Q \geq 2$, we write, as before, P "ZBC equations" for $m = -(P-1),\ldots,0$ together with additional equation for $U_k(m=1)$. We wish to establish that the functions $U(m)$ associated with the extrema (such as $k_z^{(1)}$ and $k_z^{(2)}$ in Figure 3) vanish at the extremum (denoted $\star$) while other functions distant from the extremum but degenerate with it (such as point $k_z^{(3)}$ in Figure 3) have finite amplitude $U_3(m)$ at finite $m \geq 1$.

At our disposal are P-Q distinct complex k's and Q+1 reals, the latter to be chosen from among the 2Q reals, as follows: pick one pair straddling a minimum (such as $k_z^{(1)}$ and $k_z^{(2)}$ in the Figure), and choose the remaining Q-1 reals all unpaired, from among the positive and negative k's at our disposal. The unpaired k's may be lumped together with the complex and imaginaries for the purpose of an analysis such as in part (a), and one recovers all the results (4.2)-(4.9), with a slight modification: $\sin \frac{1}{2} (k_z^{(1)} - k_z^{(2)})a$ replaces $\sin k_z a$ on the r.h.s. of Equations (4.6), (4.7) and (4.9). At an absolute maximum or minimum it is impossible to choose Q+1 reals without including at least one such pair, therefore all possible $U_k(m)$ vanish at finite m as the energy approaches its extremum. For local minima, however, such as the case illustrated here, it is possible to pick a set of k's such that no pair straddles the minimum, all remaining distinct as the energy approaches the relative extremum. For example, pick $\pm k_z^{(1)}$ and $k_z^{(3)}$, all of which remain distinct as ε is decreased through the minimum value $\varepsilon^\star$. In this case, the Vandermonde determinants all remain finite, and consequently the amplitudes and U(m) are all finite.

Thus solutions straddling a minimum (or a maximum) vanish near the surface as the minimum or maximum is approached, whereas other solutions may remain finite.

In the calculation of the surface l.d.o.s.[10] as in Equation (3.19) the extrema — local as well as absolute — will contribute little because of the vanishing of the amplitudes. Two interesting consequences follow for the quantities ρ^o in (3.19): first, ρ^o at the band edges will vanish as $|E-\epsilon_n^{min/max}|^{3/2}$ rather than as $|E-\epsilon_n^{min/max}|^{\frac{1}{2}}$, the latter being the usual behavior near the band edges for the bulk edges for the bulk d.o.s. such as (3.20). Second, the familiar spikes and cusps in the d.o.s. known as van Hove singularities,[12] caused by saddle points in the energy surface, will not appear in the l.d.o.s. (3.19) due to the vanishing of the amplitudes near the surface at these stationary points.

We have proved that the unperturbed surface l.d.o.s. $\rho^o_{n,m}$ is narrower, smoother, and (because the area must be constant) has a higher maximum at the center of the band than its bulk counterpart. As m increases, the structure characteristic of the bulk progressively reappears.

V. EXACTLY SOLUBLE SIMPLE MODEL PERTURBATION

A perturbing Hamiltonian at the surface will mix the eigenstates $U_{nk}(m)$ of the unperturbed ZBC semi-infinite solid. If the perturbation extends to P different surface planes and mixes L distinct energy bands, the solution of the scattering problem will require the diagonalization of a (P x L)-dimensional matrix, and the bound-state ("surface-state") problem requires the solution of an equally high-dimensional transcendental secular equation. In this Section we solve the

simplest non-trivial model of a perturbing Hamiltonian H_s which is diagonal in the Wannier representation ($L = 1$) and confined to the first $m=1$ surface plane (hence $P = 1$). It goes without saying that H_s has the translational symmetry of the lattice in the x-y plane; the complications which result from breaking this symmetry, as in the example of a single ad-atom, are deferred to a subsequent paper.

The matrix elements of H_s in the terminated-solid representation are:

$$(H_s)_{nk,n'k'} \equiv \int d_3r\ \Psi^*_{nk}(\vec{r})H_s(\vec{r})\Psi_{n'k'}(\vec{r}) \tag{5.1}$$

in which

$$\psi_{nk} = \psi_{\tau k_z}(\vec{r}) = (N_{x,y})^{-\frac{1}{2}} \sum_{\vec{R}_j} e^{i(k_xX_j + k_yY_j)} U_{nk}(m_j)\phi_n(\vec{r}-\vec{R}_j) \tag{5.2}$$

and $Z_j = m_ja$. For H_s diagonal in the representation of the Wannier functions $\phi_n(\vec{r}-\vec{R}_j)$, and localized at $m=1$, the matrix element is independent of k_z and k_z':

$$(H_s)_{nk,n'k'} = g_nU^*_{nk}(1)U_{nk'}(1)\delta_{k_x,k_x'}\delta_{k_y,k_y'}\delta_{n,n'} \tag{5.3}$$

in which

$$g_n = \int d_3r\ \phi_n^*(\vec{r}-\vec{R}_j)H_s(\vec{r})\phi_n(\vec{r}-\vec{R}_j)\delta_{m_j,1} \tag{5.4}$$

The eigenfunctions of the total Hamiltonian $H_o + H_s$, subject to ZBC, consist of scattering states and possibly bound states, also known as "surface states". We start by constructing the latter. Using the shorthand $\tau = (n,k_{||})$ we denote these:

$$\Phi_\tau = \sum_{k_z} F_\tau(k_z)\Psi_{\tau k}(\vec{r}) \tag{5.5}$$

Schrodinger's equation for the bound state E_τ is:

$$(H_o + H_s)\Phi_\tau = E_\tau\Phi_\tau \tag{5.6}$$

We make use of the knowledge that $H_o\Psi_{nk} = \epsilon_\tau(k_z)\Psi_{nk}$, with $\epsilon_\tau(k_z)$ a more suggestive notation for the band energy $\epsilon_n(\vec{k})$. (We also occasionally write $U_{\tau,k_z}(m)$ for $U_{nk}(m)$.) Multiplying Schrödinger's equation on the left by $\Psi^*_{nk}(r)$ and integrating over electron coordinates yields a set of equations for $F(k_z)$:

$$F_\tau(k_z)(E_\tau - \epsilon_\tau(k_z)) = \Delta g_n U^*_{nk}(1) \tag{5.7}$$

where

$$\Delta = \sum_{k'_z} U_{nk'}(1)F_\tau(k'_z)$$

Self-consistency requires either $\Delta = 0$, in which case Φ_τ is not normalizable and thus does not exist, or $\Delta \neq 0$, in which case Δ may be factored out to obtain:

$$1 = -g_n S_\tau(E_\tau) \tag{5.8}$$

where

$$S_\tau(E) \equiv \sum_{k'_z} \frac{|U_{nk'}(1)|^2}{\epsilon_\tau(k_z')-E} = \frac{a}{2\pi}\int_{-\pi/a}^{\pi/a} dk_z \frac{N_z|U_{\tau,k_z}(1)|^2}{\epsilon_\tau(k_z) - E} \tag{5.9}$$

Because $|U|^2 \propto N_z^{-1}$, $S_\tau(E)$ is a quantity $O(1)$ in the thermodynamic limit. It is also, in general, complex, when E overlaps the continuum of $\epsilon_\tau(k_z)$ between the relative band extrema. Thus for (5.8) to be

satisfied, E_τ must lie outside each continuum. This requirement does not preclude the bound state E_τ at one value of τ overlapping the continuum $\epsilon_{\tau'}(k_z)$ for a different τ', as H_s does not mix different τ's.

For g_n positive (repulsive perturbation) E_τ must lie above $\epsilon_\tau(k_{z\ max}) \equiv \epsilon_n^{max}(k_\|)$. But $(-S_\tau(E))$ achieves its maximum at $E = \epsilon_\tau(k_{z\ max})$, therefore by equations (5.8) there can be no surface state of a repulsive potential unless g_n exceeds the critical value

$$g_n \geq [S_\tau(\epsilon_\tau(k_{z\ max}))]^{-1} \approx |\alpha_{zz}^{max}|/2a^2 \qquad (5.10)$$

Estimates of the numerator of S_τ from either (4.1) or (4.16) and expansion of the denominator of S_τ about $k_{z\ max}$ using $\epsilon_\tau(k_{z\ max} + q) \approx \epsilon_\tau(k_{z\ max}) - 2|\alpha_{zz}^{max}|a^{-2}\sin^2(qa/2)$ lead to the above estimate of the r.h.s. of (5.10). (See Appendix A.) Evidently the instability against bound-states is greatest at points in the 2DBZ where the curvature α_{zz} at the relative-band maximum is the smallest (i.e. at points τ where the effective-mass is the greatest).

For g_n negative (attractive perturbation) the bound surface state energy E_τ lies below $\epsilon_\tau(k_{z\ min}) \equiv \epsilon_n^{min}(k_\|)$. We find there can be no such state unless the magnitude of the attractive potential exceeds a critical value:

$$g_n \leq [S_\tau(\epsilon_\tau(k_{z\ min}))]^{-1} \approx -|\alpha_{zz}^{min}|/2a^2 \qquad (5.11)$$

with the estimate of the r.h.s. being obtained in a similar way to (5.10). (Note that S is negative in the above.) Where $\alpha_{zz} = 0$ any arbitrary weak perturbation will cause a surface state to "peel off"; see Figures 1, 2 and 4.

If a bound state exists, it must be normalized. One readily finds for the normalized amplitudes:

$$F_\tau(k_z) = \frac{g_n U^{*}_{nk}(1)}{E_\tau - \epsilon_\tau(k_z)} \left[\partial E_\tau / \partial g_n\right]^{\frac{1}{2}} \tag{5.12}$$

in which the derivative is to be taken using (5.8).

Note that g_n may be sufficiently large to cause surface states for some $k_{||}$ but not for others. While in the n.n.s.c. (100)-band-structure equations (5.9) and (5.10) are insensitive to $k_{||}$, in the (011) orientation of the same band structure, the same α_{zz} at $k_y a_o \approx \pm \pi 2^{-\frac{1}{2}}$ results in instabilities shown in Figure 4.

Scattering theory gives us the form of the remaining states:

$$\Phi_{\tau,k_z} = \Psi_{nk}(\vec{r}) + \sum_{k' \neq k} \delta_{k_x,k_x'} \delta_{k_y,k_y} L_{k,k'} \Psi_{nk'}(\vec{r}) \tag{5.13}$$

As in the analysis of the bound state, we write Schrödinger's equation for this eigenfunction:

$$(H_o + H_s)\Phi_{\tau,k_z} = E_\tau(k_z)\Phi_{\tau,k_z} \tag{5.14}$$

or, more explicitly:

$$\begin{aligned} &\epsilon(k_z)\Psi_{k_z}(\vec{r}) + \sum_{k_z' \neq k_z} L_{k_z,k_z'} \epsilon(k_z')\Psi_{k_z'}(\vec{r}) \\ &\quad + g_n U_{k_z}(1) \sum_{k_z'} U^{*}_{k_z'}(1)\Psi_{k_z'}(\vec{r}) + g_n \sum_{k_z'} \sum_{k_z'' \neq k_z} L_{k_z k_z''} U_{k_z''}(1) U^{*}_{k_z'}(1) \Psi_{k_z'}(\vec{r}) \\ &= E(k_z)\left[\Psi_{k_z}(\vec{r}) + \sum_{k_z' \neq k_z} L_{k_z,k_z'} \Psi_{k_z'}(\vec{r})\right] \end{aligned} \tag{5.15}$$

and equate coefficients of the Ψ_k's (the subscripted τ and $k_{\|}$ omitted for typographical clarity). Starting with k_z we find:

$$E_\tau(k_z) = \epsilon_\tau(k_z) + g_n[|U_{\tau k_z}(1)|^2 + U^*_{\tau k_z}(1) \sum_{k_z'' \neq k_z} L_{k_z k_z''} U_{\tau k_z''}(1)]$$

$$= \epsilon_\tau(k_z) + g_n O(1/N_z) \qquad (5.16)$$

writing $U_{\tau k_z}$ for U_{nk}. For most purposes $E(k_z)$ can be replaced by $\epsilon(k_z)$ which it interlaces. However, if it is desired to calculate the correction, the above may be used to do so with the aid of the cofficients $L_{k,k'}$. These we now obtain by equating coefficients of $\Psi_{k_z'}(k_z' \neq k_z)$:

$$E(k_z)L_{k,k'} = \epsilon(k_z')L_{k,k'} + g_n U_{k_z}(1)U^*_{k_z'}(1) + g_n \Gamma_k U^*_{k_z'}(1)$$

where

$$\Gamma_k = \sum_{k_z'' \neq k_z} (L_{k,k''} U_{k_z''}(1)) \qquad (5.17)$$

These are readily solved, to obtain:

$$L_{k_z,k_z'} = \frac{g_n U^*_{\tau,k_z'}(1) U_{\tau,k_z}(1)}{E_\tau(k_z) - \epsilon_\tau(k_z')} [1 + g_n S_\tau(E_\tau(k_z))]^{-1} \qquad (5.18)$$

For most applications $E_\tau(k_z)$ can be replaced by its thermodynamic limiting value $\epsilon_\tau(k_z)$ in this formula. Note that the denominator $(1 + g_n S)$ cannot vanish, as S is complex in the range of interest owing to the branch cut along the continuum of the relative band. However, if there is a bound state <u>outside</u> the continuum then at a symmetric point

just <u>inside</u> the continuum one expects $(1 + g_n \mathrm{Re}(S)) = 0$, so that the amplitudes $L_{k,k'}$ are strongly enhanced there, limited only by the non-vanishing of the imaginary part Im(S). This is a <u>resonance</u>, conjugate to the bound state; see point (1) in Figure 4a.

For convenience, we rewrite the eigenstates explicitly in terms of the Wannier orbitals. Substitution of (5.2) into (5.5) yields:

$$\Phi_\tau = N_{x,y}^{-\frac{1}{2}} \sum_{R_j} e^{i(k_x X_j + k_y Y_j)} W_{\tau o}(m_j) \phi_n(r - R_j) \tag{5.19}$$

for the bound state, with

$$W_{\tau o}(m_j) = \sum_{k_z} F_\tau(k_z) U_{\tau,k_z}(m_j)$$

$$= \sum_{k_z} \frac{g_n U^*_{\tau,k_z}(1) U_{\tau,k_z}(m_j)}{E_\tau - \epsilon_\tau(k_z)} [\partial E_\tau / \partial g_n]^{\frac{1}{2}} \tag{5.20}$$

At $m_j = 1$ this has the simple limiting form,

$$W_{\tau o}(1) = [\partial E_\tau / \partial g_n]^{\frac{1}{2}} \tag{5.21}$$

It is possible to estimate the decay of the integral in (5.20) at $m_j > 1$ by using the results of the previous Section for its numerator and the effective-mass approximation for its denominator as in Appendix A, obtaining:

$$W_{\tau o}(m) \cong W_{\tau o}(1) \exp(-\gamma(m-1)) \tag{5.22}$$

with

$$\gamma = \ln \frac{1}{\frac{1}{2}(R+2) - [\frac{1}{2}(R+2)^2 - 1]^{\frac{1}{2}}} \quad \text{and}$$

$$R = 2a^2(\alpha_{zz}^{max/min})^{-1} \, |E_\tau - \epsilon_\tau(k_{z\ max/min})| \tag{5.23}$$

R being effectively the ratio of binding energy to band width for the bound state. This estimate does not replace a quantitative calculation based on actual band structures, but can serve as a guide.

We can now compute the surface-state contributions to the local density-of-states (l.d.o.s.), at fixed $\tau = (n,k_{||})$:

$$\rho_{\tau,m;bd.st.}(E) = |W_{\tau o}(m)|^2 \, \delta(E-E_\tau) \tag{5.24}$$

Proceeding with the scattering states in the same way, we write:

$$\Phi_{\tau,k_z} = N_{x,y}^{-\frac{1}{2}} \sum_{R_j} e^{i(k_x X_j + k_y Y_j)} W_{\tau,k_z}(m_j) \phi_n(r-R_j) \tag{5.25}$$

with

$$W_{\tau,k_z}(m) \equiv U_{\tau,k_z}(m) + \sum_{k_z'} L_{k_z,k_z'} U_{\tau,k_z'}(m) \tag{5.26}$$

according to (5.13). Using $L_{k,k'}$ given in (5.18) we obtain:

$$W_{\tau,k_z}(m) = U_{\tau,k_z}(m) - \frac{g_n U_{\tau,k_z}(1)}{1+g_n S_\tau(\epsilon)} \sum_{k_z'} \frac{U^*_{\tau,k_z'}(1) U_{\tau,k_z'}(m)}{\epsilon' - \epsilon} \tag{5.27}$$

abbreviating $\epsilon_{\tau}(k_z)$ by ϵ. This simplifies at $m=1$:

$$W_{\tau,k_z}(1) = U_{\tau,k_z}(1)/(1 + g_n S_{\tau}(\epsilon)) \tag{5.28}$$

Again the m-dependent integral can be estimated for $m > 1$ (see Appendix A) resulting in the following estimate of (5.27):

$$W_{\tau,k_z}(m) \cong (2/N_z)^{\frac{1}{2}} \sin\left((k_z - k_{zmin/max})ma - \theta_k\right) \tag{5.29}$$

The scattering-state (i.e. bulk-) contributions to the l.d.o.s. are:

$$\rho_{\tau,m,sc.st.}(E) = \sum_{k_z} |W_{\tau,k_z}(m)|^2 \, \delta(E - \epsilon_{\tau}(k_z)) \tag{5.30}$$

(5.24) and (5.30) have been obtained at fixed $\tau = (n,k)$. The results are sketched in Figure 4b. The following sum rule follows from completeness:

$$1 = \int_{-\infty}^{+\infty} dE \{\rho_{\tau,m,bd.st}(E) + \rho_{\tau,m,sc.st.}(E)\} \tag{5.31}$$

We also define the total band l.d.o.s. by the sum over $k_{||}$:

$$\rho_{n,m,bd.st.}(E) = N_{x,y}^{-\frac{1}{2}} \sum_{k_{||}} \rho_{\tau,m,bd.st.}(E)$$

and

$$\rho_{n,m,sc.st.}(E) = N_{x,y}^{-1} \sum_{k_{||}} \rho_{\tau,m,sc.st.}(E) \tag{5.32}$$

as sketched in Figure 4. These ρ's also satisfy the sum rule (5.31). In the limit $g_n = 0$ the bound-state d.o.s. all vanish and the scattering-state d.o.s. become equal to the unperturbed functions defined in (3.19) and (3.20), both of which have unit area. In the opposite limit (large $|g_n|$) it is possible for the scattering contributions to become small and the area sum rule to be saturated by the bound-state contributions — this is the interesting limit of complete "surface bands".

VI. VACUUM STATES

For the sake of completeness, the set of eigenstates for the solid must be augmented by a set of states for the outside, the "vacuum states". So far, we have used only the Wannier functions of atoms within the solid as our basis. Let us now test their completeness:

$$\sum_{\substack{R_j (m_j > 0) \\ \text{all } n}} \phi_n^*(r'-R_j)\phi_n(r-R_j) = \Delta_+(r',r) \tag{6.1}$$

Deep in the bulk (z or z' > Pa, with P the cut-off integer introduced in a preceeding section) we have the limiting behavior:

$$\lim_{z,z' \to \infty} \Delta_+(r',r) \to \delta(r'-r) \tag{6.2}$$

i.e., the set of Wannier functions of the solid become complete for the description of arbitrary functions deep in the bulk. But near the surface, the function $\Delta_+(r',r)$ differs — perhaps seriously so, for some applications — from the required delta function. Obviously, vacuum states are required for the description of electrons outside the solid. Less obviously, they are also required to augment the states

of the solid for a good description of the surface region, both internal and external to the solid.

Evidently, one can re-introduce the missing functions in a manner analogous to those which were retained, and define:

$$\sum_{\substack{R_j (m_j \leq 0) \\ \text{all } n}} \phi_n^*(r'-R_j)\phi_n(r-R_j) \equiv \Delta_-(r',r) \tag{6.3}$$

Now,

$$\Delta_+(r',r) + \Delta_-(r',r) = \delta(r'-r) \tag{6.4}$$

everywhere. But as the contribution of Δ_- becomes vanishingly small when r' or r are deep in the bulk, so does the contribution of Δ_+ become negligible at a comparable distance outside the solid. It is thus mainly within the surface region that both sets of states must be used, and we see that the surface region possibly has greater structure than might have been anticipated; see Figure 5.

Well outside the solid, the ϕ_n's are badly suited for describing free electron motion. The perturbances due to the presence of the solid do not manifest themselves until the surface region is reached. Thus for the vacuum states, plane waves are required, at least asymptotically. We wish to construct plane waves using the ϕ_n's, so as to ensure the orthogonality to the states inside the solid. The simplest procedure consists of defining the following basis functions:

$$\zeta_k(r) \equiv C_k^{-\frac{1}{2}} \int d_3r' \, e^{ik\cdot r'} \Delta_-(r',r) \tag{6.5}$$

in which the normalization constant is:

$$C_k = \int d_3 r' \int d_3 r'' \, e^{ik\cdot(r'-r'')} \Delta_-(r',r'') \tag{6.6a}$$

In the limit that the volume of the vacuum region $\Omega \equiv (L_z N_x a_x N_y a_y)$ is infinitely greater than the volume of the surface region $O(Pa_z N_x a_x N_y a_y)$ most of the contribution to this integral comes from the asymptotic region where $\Delta_- = \delta(r'-r'')$ and therefore,

$$C_k = L_z N_x a_x N_y a_y = \Omega, \tag{6.6b}$$

and is independent of k.

Variationally, the energy of each state $\zeta_k(r)$ is:

$$E_k = \int d_3 r \; \zeta_k^*(r)\left(-\frac{\hbar^2}{2m}\nabla^2 + H_s(r)\right)\zeta_k(r) \tag{6.7a}$$

Because any perturbations H_s are limited to the neighborhood of the surface, and because ζ_k is essentially a plane wave (except in that same neighborhood) we have:

$$E_k = \hbar^2 k^2/2m + O(Pa_z/L_z) \tag{6.7b}$$

Although normalized, orthogonal to the states of the solid, and practically "sharp" in energy, the states ζ_k suffer from two important defects. The first concerns their lack of mutual orthogonality. This is serious, for it would prevent us from remedying the second: the fact that they are not exact eigenstates of the total Hamiltonian, i.e., that a perturbation H_s, existing near the surface, may mix them with one another and with the states of the solid. The application of scattering theory, the calculation of matrix elements such as are required in the

photoelectric effect and other surface-related dynamical phenomena, problems in chemisorption, etc., all require the use of an orthonormal basis.

For this purpose, Schmidt's orthogonalization procedure fails. This becomes immediately apparent as one attempts to use it. One starts with a function, say ζ_o, orthogonalizes the remaining functions to it ($\zeta_k \rightarrow \zeta_k - (\zeta_o, \zeta_k)\zeta_o$), normalizes the new ζ_k's, singles out the next function — say ζ_{k_1} — and repeats the process, iterating until the set of functions is exhausted. As this procedure is in the nature of an algorithm, it would appear ideal for computer applications; however, suppose we wish to determine the new, properly orthonormalized, function replacing ζ_k at a finite energy E_k above the bottom of the continuum. This function can only be reached after an infinite number of steps! (This is not an unimportant objection, because it is not remedied by box-quantization, although this reduces the number of steps to a large, but finite number of steps. It is not at all clear that the orthonormal functions approach limit functions as the size of the box is allowed to approach infinity.) One desires a formulation capable of yielding the states at an energy E_k without the need for calculating all the states belonging to $E_k' < E_k$ first. A second objection is that the Schmidt procedure is totally arbitrary. If, instead of starting with ζ_o, we had started at some arbitrary ζ_k, the function ζ_o which this procedure ultimately would yield, would differ greatly from the initial ζ_o; and indeed, every function obtained in this way would differ from those obtained by the initial prescription.

For our applications, we have therefore developed a procedure which is specifically tailored to continuum states.[1] The Schmidt procedure may continue to be used for any finite subset of, e.g., localized states, while the procedure we now outline will be both practical and well suited for the continuum states. The arbitrariness in the results we have noted will be parametrized by a set of quantities ϵ_0, $\epsilon_1,\ldots$, which we can choose arbitrarily.

We first make use of the obvious identities:

$$\Delta^*_\pm(r',r) = \Delta_\pm(r,r') \text{ and } \int d_3r''\, \Delta^*_\pm(r,r'')\Delta_\pm(r',r'') = \Delta_\pm(r',r) \qquad (6.8)$$

together with (6.4) to obtain the overlaps in the form:

$$\int d_3r\; \zeta^*_k(k)\zeta_{k'}(r) = \delta_{k,k'} - (1/L_z) \sum_{\substack{R_j(m_j>0)\\ \text{all } n}} F^*_{n,R_j}(k')F_{n,R_j}(k) \qquad (6.9)$$

with:

$$F_{n,R_j}(k) \equiv (L_z/\Omega)^{\frac{1}{2}} \int d_3r\; e^{-ik\cdot r}\, \phi_n(r-R_j), \qquad (6.10)$$

obtaining (6.5) in the form:

$$\zeta_k(r) = (1/\Omega)^{\frac{1}{2}} \{e^{ik\cdot r} - (\Omega/L_z)^{\frac{1}{2}} \sum_{\substack{R_j(m_j>0)\\ n}} F^*_{n,R_j}(k)\phi_n(r-R_j)\} \qquad (6.11)$$

The vacuum region outside the solid is the given volume $\Omega = L_zN_xa_xN_ya_y$, with L_z ultimately allowed to become infinite. Thus any finite length such as Pa_z may be neglected by comparison. The arguments

given following equations (6.5) and (6.6) indicated that the wave function (6.11) vanishes for $z > Pa_z$, thus only terms with $m_j \leq P$ need be retained for $L_z < z < Pa$.

We now start our procedure, by a simple step, in which only a single function $\phi_n(r)$ is involved.

Instead of (6.11) let us first assume that the basis set consists of $\zeta_k^{(0)}$ defined as follows:

$$\zeta_k^{(0)}(r) = (1/\Omega)^{\frac{1}{2}} \{e^{ik \cdot r} - F_k^{(0)*} \phi_o(r)\} \tag{6.12}$$

in which the $\phi(r)$ to which all ζ_k's are orthogonal is normalized and

$$F_k^{(0)} \equiv \int e^{-ik \cdot r} \phi_o(r) d_3 r, \text{ which is } O(1). \tag{6.13}$$

The following results may be proved using completeness of the $e^{ik \cdot r}$:

$$(1/\Omega) \sum_k |F_k^{(0)}|^2 = 1$$

$$(1/\Omega) \sum_k F_k^{(0)} e^{ik \cdot r} = \phi_o(r)$$

$$(1/\Omega) \sum_k F_k^{(0)} \zeta_k^{(0)}(r) = 0 \tag{6.14}$$

Our procedure is to construct a pseudo-Hamiltonian in the space of the functions (6.14) having $\phi(r)$ as its ground state; the excited states will then be the desired orthonormal set. So let us define the matrix elements of $H = H_o + H'$, in the plane wave representation:

$$H_{k,k'} = \frac{\hbar^2 k^2}{2m} \delta_{k,k'} - \frac{\lambda}{\Omega} G_k^* G_{k'} \tag{6.15}$$

with λ an adjustable parameter and G_k to be determined. Evidently, the

bound state (ground state) takes the form

$$\psi_o = (1/\Omega) \sum_k g_k \, e^{ik\cdot r} \tag{6.16}$$

and has energy $-\epsilon_o$. Schrodinger's equation for the ϵ_o and g_k leads to:

$$\left(\frac{\hbar^2 k^2}{2m} + \epsilon_o\right) g_k = \frac{\lambda}{\Omega} G_k^\star \sum_{k'} G_{k'} g_{k'} \tag{6.17}$$

The non-trivial solution of these equations requires:

$$\lambda^{-1} = (1/\Omega) \sum_k \frac{|G_k|^2}{\frac{\hbar^2 k^2}{2m} + \epsilon_o} \tag{6.18}$$

and thus the normalized solution is:

$$g_k = \frac{G_k^\star}{\frac{\hbar^2 k^2}{2m} + \epsilon_o} \left[\frac{1}{\Omega} \sum_k \frac{|G_k|^2}{\left(\frac{\hbar^2 k^2}{2m} + \epsilon_o\right)^2}\right]^{-\frac{1}{2}} \tag{6.19}$$

Comparison of (6.16) and (6.14) indicates that we need $g_k \propto F_k^{(0)}$, the overlap function defined in (6.13). Thus, we choose:

$$G_k^\star = F_k^{(0)} (\hbar^2 k^2/2m + \epsilon_o) \tag{6.20}$$

which, together with (6.18) determines H'. <u>Note that ϵ_o remains arbitrary</u>; it in fact represents the extent of arbitrariness in the procedure, and parametrizes the infinite number of possible orthonormal sets which can be constructed having the desired properties. Using any convenient criterion to determine $\epsilon_o \geq 0$, or arbitrarily setting it

zero, we can construct the corresponding "perturbation pseudo-Hamiltonian":

$$H'_{k,k'} = -\frac{1}{\Omega} \frac{(F_k^{(0)*})(\hbar^2k^2/2m + \epsilon_o)(F_{k'}^{(0)})(\hbar^2k'^2/2m + \epsilon_o)}{(1/\Omega)\sum_k |F_k^{(0)}|^2(\hbar^2k^2/2m + \epsilon_o)} \quad (6.21)$$

The next step is to construct the scattering states of $H_o + H'$. They are:

$$\psi_k(r)\ (1/\Omega)^{\frac{1}{2}}, \text{ where}$$

$$\psi_k(r) = \{e^{ik\cdot r} + (1/\Omega)\sum_{k'} L_{k,k'}^{(0)}\, e^{ik'\cdot r}\} \quad (6.22)$$

We have already solved the scattering problem for a separable H' in the preceeding section. The results are:

$$L_{k,k'}^{(0)} = \frac{F_k^{(0)*}(\hbar^2k^2/2m + \epsilon_o)F_{k'}^{(0)}(\hbar^2k^2/2m + \epsilon_o)}{\frac{\hbar^2}{2m}(k^2 - k'^2)} \cdot \frac{1}{D_k^{(0)}} \quad (6.23)$$

where D_k is complex:

$$D_k^{(0)} = (\frac{\hbar^2k^2}{2m} + \epsilon_o)\frac{1}{\Omega}\sum_{k''} |F_{k''}^{(0)}|^2 \frac{\hbar^2k''^2/2m + \epsilon_o}{\frac{\hbar^2}{2m}(k''^2 - k^2)} \quad (6.24)$$

This specifies the orthonormal set of wave functions (6.22), positive-energy eigenfunctions of the synthetic Hamiltonian (6.15).

To further orthogonalize this infinite set of functions to a given $\phi_1(r)$, we merely iterate the procedure, replacing exp(ik·r) in

(6.12) - (6.16) by $\psi_k(r)$, i.e.

$$\zeta_k^{(1)}(r) \equiv (1/\Omega)^{\frac{1}{2}} \{\psi_k(r) - F_k^{(1)^*} \emptyset_1(r)\} \tag{6.25}$$

and

$$F_k^{(1)} \equiv \int \psi_k^*(r) \emptyset_1(r) d_3 r$$

$$= \int \{e^{-ik\cdot r} + (1/\Omega) \sum_{k'} L_{k,k'}^{(0)^*} e^{-ik'\cdot r}\} \emptyset_1(r) d_3 r \tag{6.26}$$

differs from (6.10), it should be noted. If, as we shall assume $\emptyset_1$ is orthogonal to $\emptyset_o$ and is normalized as well,[13] then:

$$(1/\Omega)\sum_k |F_k^{(1)}|^2 = \frac{1}{\Omega} \sum_k \int d_3 r d_3 r' \; \psi_k^*(r') \psi_k(r) \emptyset_1^*(r) \emptyset_1(r')$$

$$= \int d_3 r d_3 r' [\frac{1}{\Omega} \sum_k \psi_k^*(r') \psi_k(r) + \emptyset_o^*(r') \emptyset_o(r)] \emptyset_1^*(r) \emptyset_1(r')$$

$$= \int d_3 r d_3 r' \; \delta(r-r') \emptyset_1^*(r) \emptyset_1(r')$$

$$= 1 \tag{6.27}$$

in which we have added a term which vanishes (due to the orthogonality of $\emptyset_o$ and $\emptyset_1$) and used completeness. This proves the equivalent of the first line of (6.14). To prove the analogue of the second line, we also add a zero term containing $\emptyset_o$ and use completeness, obtaining:

$$(1/\Omega) \sum_k F_k^{(1)} \psi_k(r) = \emptyset_1(r) \tag{6.28}$$

and finally, orthogonality of the new functions to the $\emptyset_1$ ensures that:

$$(1/\Omega) \sum_k F_k^{(1)} \zeta_k^{(1)}(r) = 0 \tag{6.29}$$

The matrix elements of the new pseudo-Hamiltonian in the representation of the $\psi_k(r)\,\Omega^{-\frac{1}{2}}$ are:

$$H_{k,k'} = \epsilon_k \delta_{k,k'} - \frac{\lambda}{\Omega} G_k^* G_{k'} \tag{6.30}$$

where $\epsilon_k = \hbar^2 k^2/2m + O(1/\Omega)$, and λ and G_k for the separable interaction must be determined, as previously, so as to make $\emptyset_1$ the unique bound state, and the remaining states ipso-facto the orthonormal set of positive energy states orthogonal to $\emptyset_1$. By strict analogy, one obtains that these states take the form:

$$\psi_k^{(2)}(r)\Omega^{-\frac{1}{2}} \quad \text{for} \quad -L_z < z < Pa_z,$$

where

$$\psi_k^{(2)}(r) = \psi_k(r) + (1/\Omega) \sum_{k'} L^{(1)}_{k,k'}\, \psi_{k'}(r) \tag{6.31}$$

in which

$$L^{(1)}_{k,k'} = \frac{F_k^{(1)^*}(\epsilon_k + \epsilon_1)\, F_{k'}^{(1)}(\epsilon_{k'} + \epsilon_1)}{\epsilon_k - \epsilon_{k'}} \cdot \frac{1}{D_k^{(1)}} \tag{6.32}$$

where

$$D_k^{(1)} = (\epsilon_k + \epsilon_1)\, \frac{1}{\Omega} \sum_{k''} |F_{k''}^{(1)}|^2\, \frac{(\epsilon_{k''} + \epsilon_1)}{(\epsilon_{k''} - \epsilon_k)} \tag{6.33}$$

In the iterative scheme, $\exp(ik\cdot r)$ played the role of $\psi_k^{(o)}$, ψ_k of $\psi_k^{(1)}(r)$, etc. We now iterate any finite number of steps, and re-express the results in terms of the original plane waves:

$$\psi_k^{(n)} = e^{ik\cdot r} + (1/\Omega) \sum_{k'} M^{(n)}_{k,k'}\, e^{ik'\cdot r} \tag{6.34}$$

in which the functions, $\psi_k^{(n)} \Omega^{-\frac{1}{2}}$, form the desired orthonormal set orthogonal to the set of $\emptyset_o, \emptyset_1, \ldots, \emptyset_{n-1}$. The recursion relations are:[14]

$$M_{kk'}^{(n)} = L_{kk'}^{(n-1)} + M_{kk'}^{(n-1)} + \frac{1}{\Omega} \sum_{k''} L_{kk''}^{(n-1)} M_{k''k'}^{(n-1)} \tag{6.35}$$

$$L_{kk'}^{(n)} = \frac{F_k^{(n)^*} (\epsilon_k + \epsilon_n) F_{k'}^{(n)} (\epsilon_{k'} + \epsilon_n)}{\epsilon_k - \epsilon_{k'}} \frac{1}{D_k^{(n)}} \tag{6.36}$$

$$D_k^{(n)} = (\epsilon_k + \epsilon_n) \frac{1}{\Omega} \sum_{k''} |F_{k''}^{(n)}|^2 \frac{(\epsilon_{k''} + \epsilon_n)}{(\epsilon_{k''} - \epsilon_k)} \tag{6.37}$$

and

$$F_k^{(n)} = \int \psi_k^{(n)^*} \emptyset_n d_3 r \tag{6.38}$$

The ϵ_n's are arbitrary but should be chosen O(work function) for realism. The "energies" are $\epsilon_k = \hbar^2 k^2/2m$. While for $n \to \infty$, the normalization $\Omega^{-\frac{1}{2}}$, the energies ϵ_k, and the procedure outlined above all have to be modified for practical purposes, we orthogonalize to a finite number of planes P in the solid, belonging to L bands, so n is in fact finite and $\leq L \times P$.

VII. MORE GENERAL SURFACE PERTURBATIONS

The model perturbation treated in Section V was exactly soluble but hardly realistic; effects of a surface can be expected to extend well beyond the first surface plane. Although in scattering theory, few problems are soluble in closed form,[15] we have fortunately found two forms of surface perturbations, both extending to P surface layers and involving a number L of different bands, for which a complete and

and explicit solution can be constructed in a finite number P x L steps. They are:

(a) H_s diagonal in Wannier function representation, i.e. having band-diagonal matrix elements:

$$(H_s)_{nk,n'k'} = \sum_{m=1}^{P} g_n(m)\delta_{n,n'}\delta_{k_{||},k'_{||}}U^*_{nk}(m)U_{n'k'}(m) \qquad (7.1)$$

in the terminated-solid (ZBC) representation. This H_s may be interpreted as merely a local perturbing potential acting separately on each plane and for each band energy. It nevertheless represents a formidable problem, the solution of which requires special methods outlined below.

If the perturbing potential varies rapidly across an individual surface plane, it perforce has interband matrix elements. This feature is lacking in model (a). But introducing it can easily involve us in huge matrix diagonalization which we wish to avoid. Still, one would like to be able to study the effects of interband transitions on such important properties as recombination of electrons and holes at the surface on semiconductors, optical selection rules at the surface, etc. Therefore we develop a second model in which:

(b) H_s mixes different bands but in a "separable" manner. The matrix elements in the ZBC representation are:

$$(H_s)_{nk,n'k'} = \sum_{m=1}^{P} \gamma_n(m)\gamma_{n'}(m)\delta_{k_{||},k_{||}'}U^*_{nk}(m)U_{n'k'}(m) \qquad (7.2)$$

The band index n may also include the positive-energy (vacuum) states, provided a suitable high-energy cut-off is introduced compatible with the maximum slope of H_s across a plane.

Although the two soluble models given above represent entirely different physical situations, they may be solved by a common procedure. First, note that $\gamma_n^2(m)$ corresponds to $g_n(m)$. Thus for repulsive potentials we take $\gamma_n(m)$ real, for attractive ones pure imaginary. It is also convenient to define an operation "$\sim$": $\tilde{f} = f^*$ for all complex functions except $\gamma_n(m)$, for which $\tilde{\gamma} = \gamma^*$ if γ is real, but $\tilde{\gamma} = -\gamma^*$ if γ is imaginary, those being the only possibilities.

We now show that (a) is just a special case of (b). For with $g_n(m) = \gamma_n^2(m)$ and n limited to a single value, equations (7.2) and (7.1) coincide. By solving for each band n individually, (b) may be used to solve (a). Thus, it suffices to study (b), the more general of the two models.

Our procedure now is related to that of the preceeding Section. Instead of diagonalizing $H_o + \sum_m H_s(m)$ all at once, we first diagonalize $H_o + H_s(1)$, re-express the remaining $H_s(m)$ in the new eigenfunctions, diagonalize $(H_o + H_s(1)) + H_s(2)$, repeat the performance, etc. After P steps the total Hamiltonian is diagonal, and all the eigenfunctions and eigenvalues — both continuous and discrete — are known in the presence of the complete perturbation.

Let us work out the mth step. We suppose that $H^{(m-1)} = [H_o + \sum_{m'=1}^{m-1} H_s(m')]$ is already diagonalized, and now wish to work out the effects of the mth perturbation, $H_s(m)$. In the representation of the eigenstates of $H^{(m-1)}$, the perturbation has the matrix structure:

$$(H_s(m))_{q,q'} = t_q(m)\tilde{t}_{q'}(m)\delta_{k_{\parallel},k'_{\parallel}}, \qquad (7.3)$$

where

$$t_q(m) = \gamma_q(m) W_q^{(m-1)}(m) \tag{7.4}$$

and q a quantum number running over the continuum states (n,k) as well and the bound states of $H^{(m-1)}$. $W_q^{(m-1)}(p)$ is the amplitude of the corresponding eigenstates on the pth plane. Evidently,

$$t_q(m) = \begin{cases} O(N_z^{-\frac{1}{2}}) \text{ for q in the continuum of } H^{(m-1)}, \\ \\ O(1) \text{ for q ranging over the bound states} \\ \text{of } H^{(m-1)} \end{cases} \tag{7.5}$$

The eigenstates of the new Hamiltonian $H^{(m)} \equiv H^{(m-1)} + H_s(m)$ are obtained by the standard methods used in previous Sections V and VI, and result in new amplitudes given by:

$$W_q^{(m)}(p) = M_q^{-\frac{1}{2}} [W_q^{(m-1)}(p) + \sum_{q'} L_{q,q'}^{(m)} W_{q'}^{(m-1)}(p)] \tag{7.6}$$

with

$$L_{q,q'}^{(m)} = \frac{\tilde{t}_{q'}(m)\, t_q(m)}{E_q - \epsilon_{q'}} [1 + \mathcal{S}_m(E_q)]^{-1} \tag{7.7}$$

and

$$\mathcal{S}_m(E) = \sum_{q''} \frac{\tilde{t}_{q''}(m) t_{q''}(m)}{\epsilon_{q''} - E} \tag{7.8}$$

The normalization constant may be guessed to be trivally $1 + O(N_z^{-1})$ for the scattering (continuum) states; in general it is:

$$M_q = 1 + t_q(m)\tilde{t}_q(m) \frac{\partial}{\partial E_q} \ln[1 + \mathcal{S}_m(E_q)] \tag{7.9}$$

and the new energies $E_q \equiv \epsilon_q^{(m)}$ are related to the previous eigenvalues $\epsilon_q \equiv \epsilon_q^{(m-1)}$ by:

$$E_q = \epsilon_q + t_q(m)\tilde{t}_q(m) [1 + \mathcal{S}_m(E_q)]^{-1} \tag{7.10}$$

Thus, by (7.5) for continuum states the energy corrections are negligible, and the new eigenvalues interlace the old. For bound states of $H^{(m-1)}$ the amplitudes t_q are O(1) and the energies are shifed by a finite fraction, and equation (7.10) is then a non-trivial transcendental equation for the new energies. Finally, new bound states can appear as a consequence of the perturbation at m. These will appear as new roots E of:

$$1 + \mathcal{S}_m(E) = 0 \tag{7.11}$$

at energies E which differ from any of the set of E_q's we have obtained (either bound- or continuum states) using (7.10). The formulas (7.6) - (7.9) for the eigenstates can still be used for the new bound states; replacing E_q by E the solution of (7.11) in these formulas, we observe that the new pole in L dominates the rest, that it appears both in the numerator and in the normalization denominator, that the terms $t_q(1 + \mathcal{S})^{-1}$ will cancel thereby, thus leaving a remainder in precisely the form of the normalized bound state derived in (5.5) and (5.12).

We proceed to the next stage. Setting,

$$t_q(m+1) = \gamma_q(m+1)W_q^{(m)}(m+1) \quad (7.12)$$

where $W_q^{(m)}(m+1)$ the amplitude (7.6) at $p = m+1$, one solves for the effects of the perturbation at $m + 1$. Equation (7.3), with $m + 1$ replacing m, now defines $H_s(m+1)$. The procedure is repeated. P such iterations following the initial condition $W_q^{(o)}(p) = U_{nk}(p)$ suffice to obtain explicitly the exact eigenstates and eigenvalue spectrum of the total perturbed Hamiltonian $H^{(P)} = H_o + H_s$, and the corresponding amplitudes $W_q^{(P)}(p)$ on the pth plane, for any $1 \leq p \leq N_z$.

VIII. CONCLUSION

In this paper we have related the band structure to the notion of the chemical bond, have shown how the breaking of bonds at a surface alters the eigenstates, and how additional perturbations can create bound states among these — identified as the surface bands. The extinction of the bulk van Hove singularities in the surface density of states is an interesting new result, which we proved rigorously for the terminated solid. The introduction of ordinary perturbing Hamiltonians affects surface amplitudes smoothly, so the disappearance of van Hove singularities is undoubtedly a completely general feature. Of course, new van Hove singularities in cases where there exist 2D band of surface states can be expected, but they are then surface-related properties. Our theory can have consequences in the interpretation of photoelectric emission. This important probe is one of several which are specific to surfaces and require a good understanding of the positive-energy (vacuum) states. For the latter we have developed an

orthonormalization program which is effective for continuum states and complements Schmidt's method for orthonormalization of discrete states. We also recapitulated our earlier criterion[1] which differs from those which are presently well known[16] for the instability against the formation of surface states against arbitrary perturbations. Now, let us turn to the <u>nature</u> of these perturbations.

Surface perturbations have no universal character and so must be decided on a case-by-case approach. The study of metals will differ from insulators, and both are different from the large variety of semiconductor terminations. Let us start with the insulators: we have seen that the eigenstates are in one-to-one correspondence with the unperturbed states, so that if we populate the valence band, including surface states associated with it, charge neutrality is ensured. However, the introduction of the surface has invariably contributed to a surface dipole moment. (Electrons have been removed, the wavefunctions of which extended into the solid, whereas the compensating ionic charges lay outside.) The potential associated with this dipole layer represents the variation in the Madelung potential — the electrostatic potential felt by the electrons at any site — at the neighborhood of the surface. The potentials and the charge imbalances they engender must be related. Poisson's equation and the Hartree approximation may achieve this relation, as may the Hartree-Fock method. Such methods have indeed been applied to surfaces, and a vast literature now exists on them.[16] Nevertheless, applications using the present methodology may have advantages over the previous ones, many of which[17] require the <u>ab-initio</u> solution of Schrödinger's differential equation and make

little or no use of the information already available from the study of the bulk.

The problems of metals are different but equally challenging. The electron states are populated up to the Fermi energy, which is determined by the bulk. But the states lying below ε_F may have small amplitudes at the surface, and thus the surface electronic charge may be deficient. We see that this would result in an attractive corrective potential, which in turn would increase the amplitude of the states at the surface, until charge neutrality is achieved. Either Hartree or Hartree-Fock methods may be used here or the more modern schemes developed by Kohn and his collaborators.[18] However, one suspects that, as in metals, a judicious application of Friedel's sum rule[19] may prove useful, decisive, or at least, a simplification. We are presently investigating this possibility as well as those several other issues: catalysis, ad-atoms, etc., which are logically related to the present work.

APPENDIX A

Near an extremum $U \sim \sin qa$ and $\epsilon(k_z) \approx \epsilon_{max/min} \mp 2|\alpha_{zz}|a^{-2}\sin^2\frac{1}{2}qa$, where $q \equiv k_z - k_{zmax/min}$. We use these expansions not just for small q, but throughout the BZ to estimate such quantities as:

$$N_z^{-1} \sum_{k_z} U^*(1)U(m)[E - \epsilon(k_z)]^{-1} \tag{A.1}$$

Thus, the above is approximately:

$$I_m = \frac{a}{\pi}\int_{-\pi/a}^{\pi/a} d\eta \sin\eta a \sin\eta am[\Delta + 2|\alpha_{zz}|a^{-2}\sin^2\tfrac{1}{2}\eta a]^{-1} \tag{A.2}$$

where, for E outside the continuum (bound state) we derive (5.22) as follows: let

$$\Delta \equiv |E - \epsilon_{max/min}| \tag{A.3}$$

be the binding energy, and R the dimensionless quantity $R = 2\Delta a^2/|\alpha_{zz}|$. With J_m defined by $I_m = 2a^2J_m/|\alpha|_{zz}$, we have:

$$J_m = \frac{1}{\pi}\int_{-\pi}^{\pi} d\eta \sin\eta \sin\eta m [R + 4\sin^2\tfrac{1}{2}\eta]^{-1}$$

$$= \frac{1}{2\pi i}\oint d\zeta(\zeta^{m-1} - \zeta^{m+1})[(R+2)\zeta - \zeta^2 - 1]^{-1} \tag{A.4}$$

where the contour is the unit circle. Evaluating the residue at:

$$\tfrac{1}{2}(R+2) - [\tfrac{1}{4}(R+2)^2 - 1]^{\frac{1}{2}} \tag{A.5}$$

we obtain:

$$J_m = J_1 \exp-(m-1)\gamma \tag{A.6}$$

where J_1 is expressible precisely in terms of $S_\tau(E)$, and γ is defined in (5.23).

For E in the continuum, we have replaced E by $\epsilon(k_z)$ and obtain:

$$J_m = J_o \, e^{iqam} \tag{A.7}$$

where J_o is real and $J_1 \propto S_\tau(\epsilon(k_z))$. Thus,

$$S_\tau(\epsilon(k_z)) \approx |S_\tau| \exp iqa \tag{A.8}$$

For the derivation of (5.29), we use

$$U \approx (2/N_z)^{\frac{1}{2}} \sin qam \tag{A.9}$$

and insert into (5.27):

$$W \approx (2/N_z)^{\frac{1}{2}} \left[\sin qma - \frac{g_n S_\tau}{1+g_n S_\tau} e^{iq(m-1)a} \sin qa\right] \tag{A.10}$$

With the use of (A.8), we obtain the desired estimate (5.29).

FOOTNOTES AND REFERENCES

1. Small parts of this work have been published: D. Mattis, Physics Letters 60A, 468 (1977), D. Mattis, "Procedure for Orthogonalizing States in the Continuum",

2. See reviews in C.G. Scott and C.E. Reed, "Surface Physics of Phosphors and Semiconductors", esp. article by R.O. Jones, "Aspects of Surface State Theory", pp. 96-143, therein, Academic Press, New York, 1975. Also, a method similar in spirit if not in detail, to the present work: V. Bortolani, V. Celli, A. Marvin and G. Santoro, "A New Method for the Determination of Surface States", pp. 714-717, in "Physics of Semiconductors", Proc. of the 13th Int'l. Conf., Rome, 1975, G.F. Fumi, Ed., North-Holland Publ. Co., New York, 1976. The tight-binding approach was pioneered by Hückel and is often given his name. A review of its applications is given by A.J. Bennett in "Some Electronic Applications of Solid Surfaces", pp. 261-278 in Critical Review in Solid State Science 4, Issue 2, 1974, Schuele and Hoffman, eds. The first specific application to surfaces of solids may be noted in a series of publications by J. Koutecky, starting with Phys. Rev. 108, 13 (1957).

3. This feature, partly conjectured in reference 1, and important for the determination of the photoelectric spectrum, criterion for bound states energy, etc., is definitively proved in the present work.

4. Scott and Reed, reference 2; also see Chapter 4 in S.R. Morrison, "The Chemical Physics of Surfaces", Plenum Press, New York, 1977, where it is shown how, particularly in ionic solids, the Madelung potentials may assume different values for different energy bands Lett. near the surface. S. Trullinger and S. Cunningham, Phys. Rev. 30 913 (1973) have shown that surface charge imbalance can dynamically cause surface reconstruction to occur.

5. Second paper of reference 1.

6. A more general term "relocation" is employed by Morrison, reference 4.

7. A description of tight binding is in most solid-state texts, such as W.A. Harrison's "Solid State Theory", McGraw-Hill Book Co., New York, 1970, pp. 151 et seq. See also Applications by Koutecky Reference 2, and J. Davenport, T. Einstein and J.R. Schrieffer, Proc. of 2nd International Conference on Solid Surfaces, Kyoto, Japan, (1974).

8. Harrison, op. cit (reference 7), Chapter II. Applications of the popular "pseudopotential method" to surfaces is reviewed by Bennett, op. cit. Reference 2. See J.A. Appelbaum and D.R. Hamann: Phys. Rev. Lett. 31, 106 (1973), 32, 225 (1974) and Revs. Mod. Phys. 48, 479 (1976).

9. A number of papers concern the numerical study of slabs of various thicknesses. As an example, the resolvent method that "works for finite crystals (in practice up to about 1000 atoms) with two surfaces parallel to a chosen plane": A. van der Avoird, et al, Phys. Rev. B10, 1230 (1974). However some authors have tackled the semi-infinite solid directly, e.g. J.A. Appelbaum and D.R. Hamann, Phys. Rev. Letts 31, 106 (1973) and 32, 225 (1974).

10. See definition and plots of l.d.o.s. (called LDS therein) in J.R. Schrieffer and P. Soven, Physics Today, April 1975 issue; pp. 24-30.

11. See: J. Koutecky, Phys. Rev. 108, 13 (1957); J. Koutecky and M. Tomasek, Surface Science 3, 333 (1965); D. Kalkstein and P. Soven, Surface Science 26, 85 (1971).

12. Well known discontinuities in the d.o.s. in the variational and electronic band structures of solids; see van Hove, Phys. Rev. 89, 1189 (1953), who first derived them on topological grounds, or any solid state text such as: J. Ziman, "Principles of Theory of Solids", Cambridge University Press, 1964.

13. If it is not, Schmidt orthogonalize and normalize them first.

14. This method of solution of the multiple-scattering problem may find other applications, as in the study of not-so-dilute impurities in solids. For another application, see Section VII herein.

15. One often has to be content with expressing the scattered states as solutions to a Fredholm integral equation, to be developed as an infinite series, or approximated variationally. In the present work, we develop only the exactly soluble cases.

16. Aside from the books cited elsewhere, a considerable bibliography is available in Physics Today, April 1975 issue. The pioneering works are by I. Pamm, Phys. Zeits. Sowj. 1, 733 (1932) and W. Shockley, Phys. Rev. 56, 317 (1939).

17. But not all — in some schemes, the surface solutions are matched to the known bulk solution at the pth plane; see Appelbaum and Hamann, op. cit.[9]

18. P. Hohenberg and W. Kohn, Phys. Rev. B136, 864 (1964) and W. Kohn and L.J. Sham, Phys. Rev. A140, 1133 (1964), applied to surfaces by N.D. Lang and W. Kohn, Phys. Rev. B1, 4555 (1970 and B3, 1215 (1971), and J. Perdew and R. Monnier, Phys. Rev. Lett. 37, 1286 (1976) and D.C. Langreth and J. Perdew, Solid State Communication 17, 1425 (1976).

19. Discussed in many texts, such as Ziman, op. cit (reference 12 above).

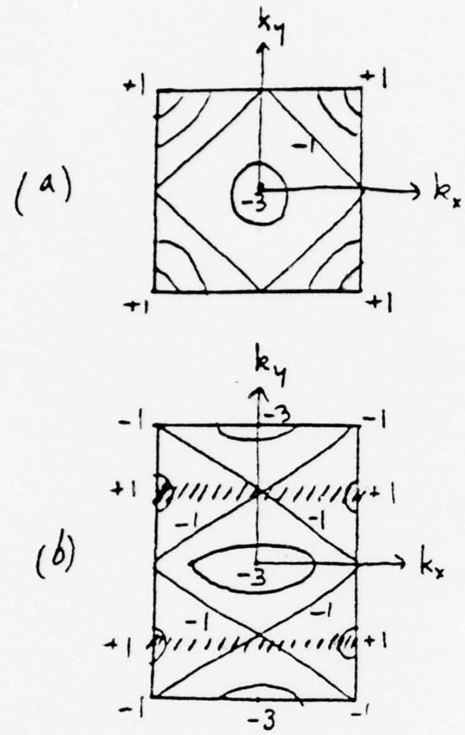

Figure 1

Contours of constant ϵ_{min} for "cubium" in the 2D B.Z.'s appropriate to:

(a) (001) surface orientation $\epsilon_{min} \propto -(\cos k_x a_o + \cos k_y a_o + 1)$

(b) (011) surface orientation $\epsilon_{min} \propto -(\cos k_x a_o + 2|\cos k_y a_o 2^{-\frac{1}{2}}|)$

(schmatically). Regions of instability against surface states are indicated by shading, and are obtained from equation (5.11) and Figure 2). These exist only for the (011) orientation and over the restricted energy range $-1 \leq \epsilon \leq +1$; the <u>full</u> band extends from -3 to +3.

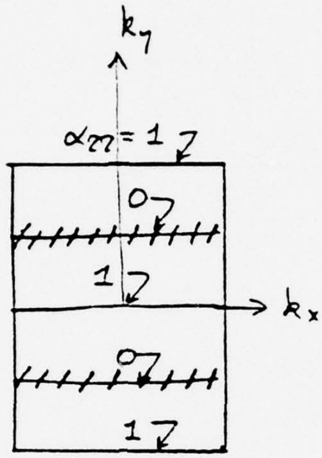

Figure 2

Loci of constant α_{zz}^{min} for (011) surface of cubium. Shaded region near $\alpha_{zz} = 0$ indicates instability against the formation of bands of surface-states for arbitrarily weak surface perturbations. Such regions of instability are generally curved, for arbitrary band structures.

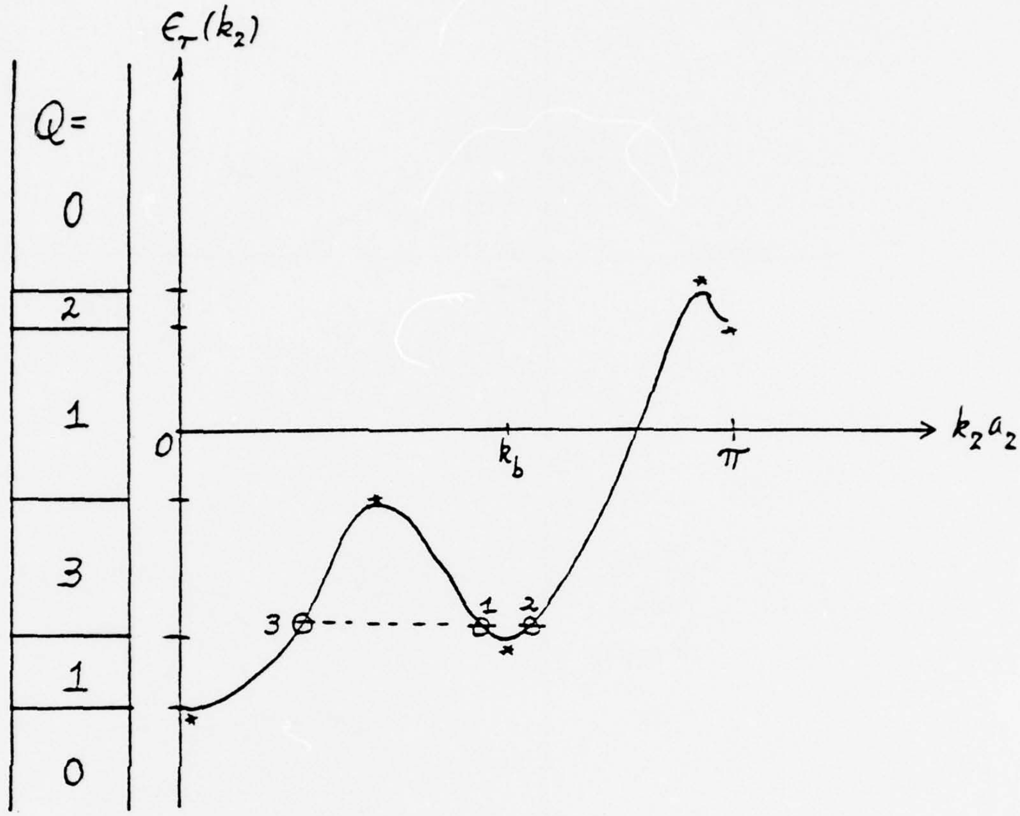

Figure 3

Typical energy contours at $\tau = (n,k_x,k_y)$. The independent variable k_z is just one of the roots of the equation $\epsilon_\tau(k_z)$ = const., where $\epsilon_\tau(k_z)$ is given in equation (3.6). The total number of positive real roots of this equation is Q, indicated on the left-hand column. In the text (Sections III and IV) it is shown that only Q of these, coupled with a number of the complex roots, survive the imposition of surface boundary conditions. It is also proved that surface amplitudes vanish at points marked (*) where Q changes discontinuously.

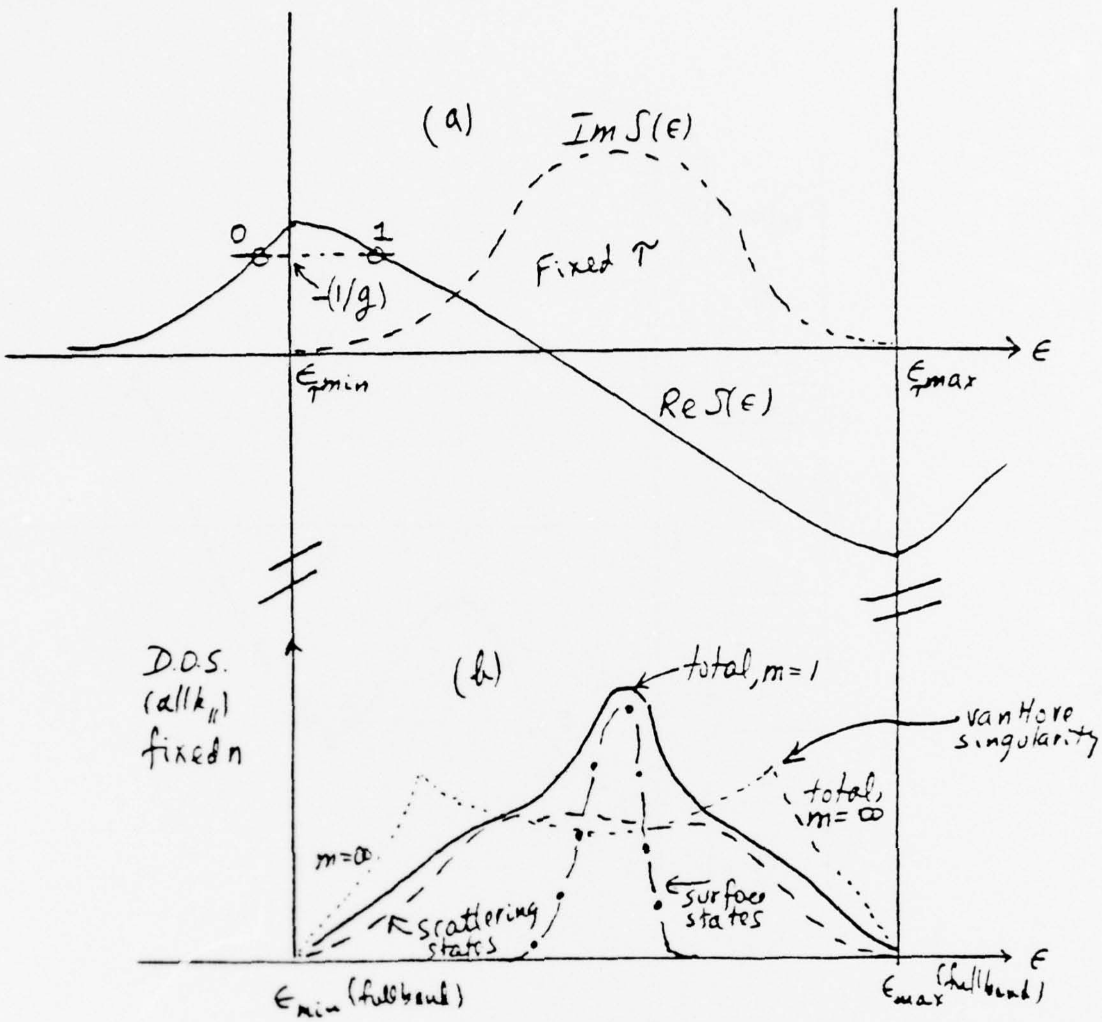

Figure 4

Bound states and their effect on the d.o.s. (schematic).

(a) Graphical construction of solution of equation (5.8) for a single band. "o" indicates the bound state, "1" the conjugate resonance, for the example of an attractive perturbation ($g < 0$).

(b) L.d.o.s. for (011) cubium, $m = 1$. For weak perturbations, as assumed in this figure, the surface (bound) states contribute only over a narrow range, indicated in the shaded regions of Figures 1 and 2. In the bulk ($m = \infty$), the surface contributions vanish, and van Hove singularities re-appear.

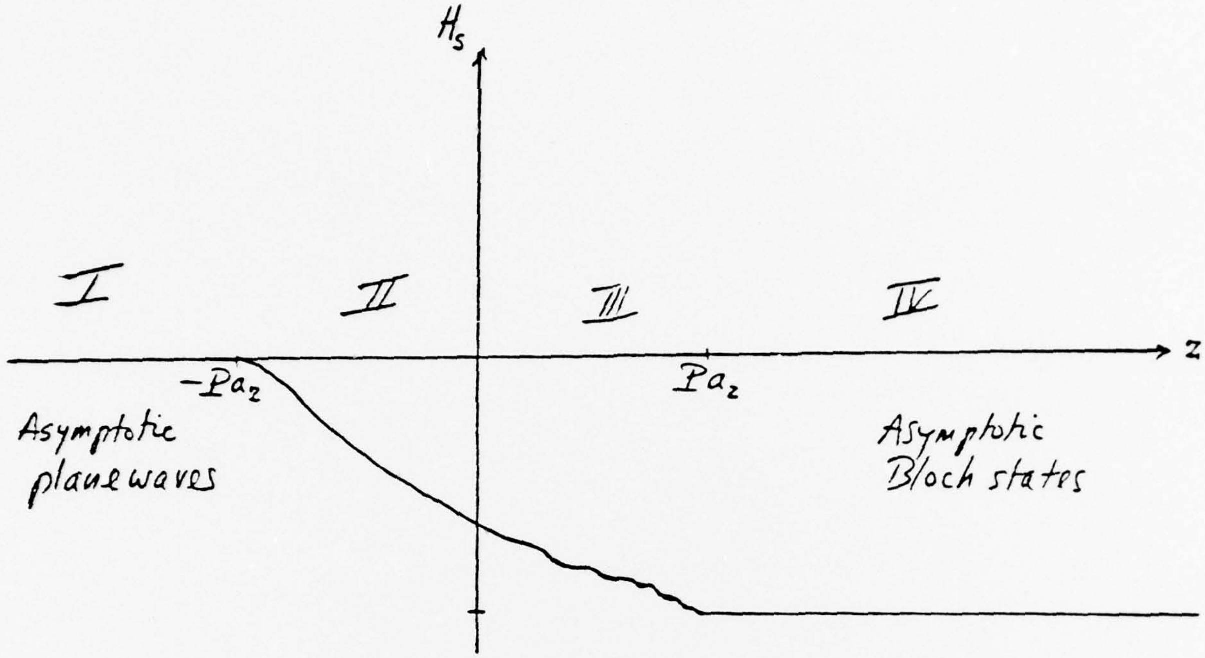

Figure 5

Principal surface regions and perturbation (schematic).

(I) Vacuum region, extending from $z = -Pa_z$ to $-L_z$ ($L_z \to \infty$). Eigenfunctions are asumptotically plane waves having positive energies $\hbar^2 k^2/2m$; a complete set.

(II) Finite-sized vacuum interface region, $O(Pa_z)$. Two sets of functions are needed for completeness.

(a) Plane waves are modified here by orthogonalization to eigenstates of the solid and to each other.
(b) Eigenstates of the solid leak into this region, to varying extent.

(III) Finite-sized solid interface region, $O(Pa_z)$, also two sets of functions are required.

(a) States of the solid are affected strongly by varying H_s.
(b) Some vacuum states leak into here.

(IV) Asymptotic solid. The eigenstates are the essentially complete set of Bloch functions; the solid has arbitrarily large length $N_z a_z$ (although $(N_z a_z / L_z) \to 0$). It should be noted that there is a distinction between the vacuum states and the positive energy states of the solid, regardless whether matrix elements connect them or not. Both sets are required for completeness. Cross-sections are $N_x a_x N_y a_y$, same in all regions.

ELECTRONIC INSTABILITY OF SURFACES OF SOLIDS*

Daniel C. MATTIS

Belfer Graduate School of Science, Yeshiva University, New York, N.Y. 10033, USA

Received 1 February 1977

We study the effects of a surface perturbation on a semi-infinite solid. The bulk band structure is found to determine whether or not the surface is unstable against the formation of electronic surface bands. A criterion, involving a single component of the bulk inverse-effective-mass tensor, is derived.

Experimental [1] and theoretical ‡1 analysis have by now established that the surfaces of solids often differ in crystallographic class [3] as well as in electronic properties ‡2 from the bulk. The fact that not all surfaces have been found to be reconstructed led me to seek a criterion, and ultimately to pose a simpler problem: under what circumstances will an infinitesimal perturbation produce a band of surface states? We shall denote this an "intrinsic instability", for while any arbitrary surface will have bound states only if the perturbing potentials at the surface are sufficiently great, one with intrinsic instability *always must* ‡3. There have been previous attempts at systematizing criteria for the existence of surface states [6]; the charge redistribution effected by their existence may additionally be responsible for the observed surface reconstruction [3]. In this paper we offer mathematical proof that a surface defining the $x-y$ plane is intrinsically unstable iff a component of the bulk inverse-mass tensor α_{zz} vanishes along a set of points or lines in the Brillouin Zone. We thus focus on the topology of the bulk band structure as a root cause of the surface instability: the exact nature and strength of the surface perturbations H_s are of secondary importance. (Our analysis does not require any particular methodology such as tight-binding, NFE, APW, etc.)

The prototype material consists of atomic planes at $Z_j = m_j a$, with $m_j = 1, 2, ..., N_z$ the plane index. The interplanar distance $\boldsymbol{a}$ is a function of the orientation of the surface relative to the crystal axes, of course. The unperturbed Hamiltonian H_0 of the terminated solid has eigenfunctions $\Psi_{nk}(\boldsymbol{r})$, in which k_x and k_y are the usual crystal momenta but k_z is restricted to the interval $0 \leqslant k_z \leqslant \pi/a$. These eigenfunctions have been shown [7] to take the form:

$$\Psi_{nk}(\boldsymbol{r}) = N_{x,y}^{-1/2} \sum_j e^{i(k_x X_j + k_y Y_j)} U_{nk}(Z_j)\phi_n(\boldsymbol{r}-\boldsymbol{R}_j)\,, \tag{1}$$

in which $\phi_n(\boldsymbol{r}-\boldsymbol{R}_j)$ is the usual Wannier function (the Fourier transform of the Bloch functions for the nth band using periodic boundary conditions). While termination of the solid preserves the translational invariance in the $x-y$ plane (witness the exponential factors) it has the effect of replacing $N_z^{-1/2}\exp(ik_z Z_j)$ by the components of a new $N_z \times N_z$ unitary matrix $U_{nk}(Z_j)$. The construction of this function and its analytical properties are elaborated in a companion paper [7]. We quote one result: suppose the eigenvalue corresponding to (1), denoted $\epsilon_n(k_x, k_y, k_z)$ has a minimum or maximum at $k_z = k_m$ (for fixed k_x, k_y and $\boldsymbol{n}$). Defining $\delta \equiv (k_z - k_m)a$, the following is of importance:

$$\lim_{\delta\to 0} |U_{nk}(m_j a)|^2 \leqslant (2/N_z) m_j^2 \delta^2\,, \tag{2}$$

for m_j any positive, finite, integer. By contrast, deep in the bulk ($m_j = +\infty$) $|U_{nk}|^2 = 1/N_z$ is a constant independent of k_z and outside the solid ($m_j = 0$ or ne-

* This research supported by a grant from the Office of Naval Research.

‡1 Some theoretical concepts are reviewed in [2]; note that they number the first atomic plane $n = 0$, whereas in the present work it is $n = 1$. See also chapters by Jones and Berg in [1].

‡2 Such as formation of energy levels bound to the neighborhood of a surface-surface states. See ref. [4] and eq. (4) of the present paper.

‡3 This is because surface perturbations must always exist. For example, the Madelung potential must deviate in the neighborhood of a surface from its bulk limiting value. See such discussions as in [5].

tive integer) $U_{nk} \equiv 0$. Thus, (2) is purely a property of the interior neighborhood of the surface.

We now introduce a perturbing Hamiltonian H_s which, for simplicity, is taken diagonal in the Wannier function representation [‡4]. It follows that its matrix structure in the basis set (1) is:

$$(H_s)_{nk,n'k'} \equiv \int \mathrm{d}_3 r\, \Psi^*_{nk}(\boldsymbol{r}) H_s(\boldsymbol{r}) \Psi_{n'k'}(\boldsymbol{r})$$

$$= \sum_p g_{n,p} U^*_{nk}(pa) U_{nk}(pa) \delta_{k_x,k_x'} \delta_{k_y,k_y'} \delta_{n,n'} \ . \quad (3)$$

By construction, H_s preserves translational invariance in the $x-y$ plane and thus conserves momenta k_x, k_y. The quantities $g_{n,p}$, the coupling-constants, parametrize the strength of the perturbation at the pth plane for the nth band.

We now study bound states of $H_0 + H_s$. They are required to be of the form:

$$\zeta_{n,k_x,k_y,s} = N_z^{-1/2} \sum_{k_z} F_{n,k_x,k_y}(k_z) \Psi_{\boldsymbol{n}\boldsymbol{k}}(\boldsymbol{r}) \ . \quad (4)$$

This is to be contrasted with bulk states, which take the form:

$$\zeta_{n,k_x,k_y,k_z} = \Psi_{\boldsymbol{nk}}(\boldsymbol{r}) + N_z^{-1} \sum_{\boldsymbol{k}'} \delta_{k_x,k_x'} \delta_{k_y,k_y'}$$

$$\times L_{\boldsymbol{nk},\boldsymbol{nk}'} \Psi_{\boldsymbol{nk}'}(\boldsymbol{r}) \quad (5)$$

in the presence of the surface perturbation H_s. The exact scattering theoretic calculation of $L_{nk,nk'}$ is given in a separate paper [‡4]. For present purposes it is quite sufficient to study the bound states (4), if any, and their energies $E_{n,k_x,k_y,s}$. As k_x, k_y and $\boldsymbol{n}$ are kept fixed, we omit these subscripts in what follows for greater topographical clarity. First, Schrödinger's equation $(H_0 + H_s)\zeta_s = E_s \zeta_s$ yields the following equation for the coefficients $F(k_z)$:

$$\sum_{k_z} F(k_z) \epsilon(k_z) \Psi_{k_z} + \sum_{p,k_z,k_z'} g_p U^*_{k_z}(p) U_{k_z'}(p) F(k_z') \Psi_{k_z}$$

$$= E_s \sum_{k_z} F(k_z) \Psi_{k_z} \ . \quad (6)$$

in which $\epsilon(k_z)$ is the bulk band structure $\epsilon_n(\boldsymbol{k})$ with superfluous indices suppressed. Equating coefficients of Ψ_{k_z}'s we readily obtain a set of linear, homogeneous equations for the $F(k_z)$'s, and a condition for a nontrivial solution to exist:

$$\mathrm{Det}\,\|\delta_{p,p'} - g_p S_{p,p'}(E_s)\| = 0 \ . \quad (7)$$

The analysis of this secular determinant is the main object of this paper. The $S_{p,p'}$ are integrals:

$$S_{p,p'}(E) = \frac{a}{\pi} \int_0^{\pi/a} \mathrm{d}k_z \frac{u^*_{k_z}(p) u_{k_z}(p')}{E - \epsilon(k_z)} \ . \quad (8)$$

We have absorbed a factor of $N_z^{1/2}$ in U_k to form the $u_k \equiv U_k N_z^{1/2}$, independent of N_z in the thermodynamic (large N) limit. For E_s within the continuum range of $\epsilon(k_z)$ the $S_{p,p'}$ are complex and therefore (7) has no solution. Thus the bound states lie outside the extrema at k_m, while quite possibly overlapping some of the bulk continuum of states for different (k_x, k_y)'s or different band index $\boldsymbol{n}$.

As we have seen in (2), the numerators in (8) vanish when k_z approaches an extremum. Therefore the $S_{p,p'}$ are ordinarily finite integrals, no matter how closely E_s approaches the continuum edge at $\epsilon(k_m)$. Now let us examine $\epsilon(k_z)$, i.e. $\epsilon_n(\boldsymbol{k})$, near an extremum:

$$\epsilon_n(\boldsymbol{k}) = \epsilon_n(k_x, k_y, k_m) + \tfrac{1}{2}\alpha_{zz}(k_x, k_y, k_m)(k_z - k_m)^2 a^2$$

$$+ \mathrm{O}((k_z - k_m)^4) \ , \quad (9)$$

where $\alpha_{zz} \equiv \partial^2 \epsilon_n(\boldsymbol{k})/\partial(ak_z)^2$ is the zz-component of the inverse-effective-mass tensor, to be evaluated at the point (k_x, k_y, k_m) in the nth band [‡5]. While normally positive at a minimum and negative at a maximum, this quantity could be zero at either, causing $S_{p,p'}$ to diverge as $|E_s - \epsilon(k_m)|^{-1/4}$ when E_s approaches $\epsilon(k_m)$. Thus, the criterion for (7) to have a solution, when the coupling constants g_p are arbitrarily small, is merely:

$$\alpha_{zz}(k_x, k_y, k_m) = 0 \ . \quad (10)$$

Wherever this is satisfied, we necessarily have surface states. This may occur at discrete points or along curves in the (k_x, k_y) plane. By continuity, at finite values

[‡4] In a more complete study of the perturbed surface, under rpeparation, we consider inter-band matrix elements, solve for scattering states and charge redistribution, density-of-states functions, etc.

[‡5] The behavior of U, eq. (2), is purely geometrical and unrelated to the magnitude of α_{zz}. Thus any divergent behavior of the $S_{p,p'}$ integrals is caused solely by the energy denominators.

of g_p a finite neighborhood of these points or curves also necessarily satisfies the condition (7) for the existence of a bound surface state.

By way of an illuminating application, consider the s.c. tight-binding band structure: $\epsilon(k) = -K(\cos k_x a + \cos k_y a + \cos k_z a)$, for a physical surface perpendicular to the (100) axis. The band extrema are at $k_m = 0$ or π/a, and α_{zz} takes on the values $+K$ at the one and $-K$ at the other, for all (k_x, k_y). Thus, the criterion eq. (10) for intrinsic instability is met nowhere in the 2D Brillouin zone. On the other hand, tilting the surface perpendicular to the (011) direction results in new lattice spacings: $a_y = a_z = a/2^{1/2}$ and $a_x = a$, and in the new coordinate system (k_z axis must always be perpendicular to the physical surface) the energy-band structure is: $\epsilon(k) = -K(\cos k_x a + 2 \cos k_y a_y \cos k_z a_z)$. The extrema are at $k_m a = 0$ or π, and eq. (10) now has nontribial solutions along the segments $k_y a_y = \pm\frac{1}{2}\pi$, for all k_x. This orientation surface is intrinsically unstable, whereas the other orientation was not, for the same material.

References

[1] See Physics Today issue on surface physics, 28 (April 1975); also, J.A. Appelbaum and D.R. Hamann, Rev. Mod. Phys. 48 (1976) 479;
C.G. Scott and C.E. Reed (eds.), Surface physics of phosphors and semiconductors (Academic Press, New York, 1975).
[2] J.R. Schrieffer and P. Soven, Physics Today 28 (1975) 24.
[3] P.J. Estrup, Physics Today 28 (1975) 33.
[4] J.A. Appelbaum and D.R. Hamann, Phys. Rev. Lett. 32 (1974) 225.
[5] A. Clark, The chemisorptive bond (Academic, New York, 1974) sect. 9.3;
J. Perdew and R. Monnier, Phys. Rev. Lett. 37 (1976) 1286.
[6] W. Shockley, Phys. Rev. 56 (1939) 1317. A mere recent criterion proposed by S.J. Gurman and J.B. Pendry, Phys. Rev. Lett. 31 (1973) 637 has been strongly disputed by L. Kleinman, Phys. Rev. B13 (1976) 4640.
[7] D.C. Mattis, The terminated solid: Exact eigenfunctions of an idealized surface, to be published.

PHYSICAL REVIEW B VOLUME 15, NUMBER 8 15 APRIL 1977

Electron states in random alloys with short-range order

Paul Bloom* and Daniel Mattis†

Belfer Graduate School of Science, Yeshiva University, New York, New York 10033

(Received 3 August 1976)

We present an accurate and economical iterative method of calculating the energy levels of a disordered or partly ordered random alloy. Results presented for one- and three-dimensional simple cubic lattices compare favorably with exact calculations. We also present the systematic effects of partial short-range order in three dimensions. A theory of the one-particle propagators is presented, and the theory of electrical conductivity is developed in the context of our new method. Our formulas satisfy the exact conservation laws.

I. INTRODUCTION

The study of electronic and vibrational spectra of disordered alloys is currently one of the principal concerns of solid-state physics,[1] stimulated by the outstanding successes of the coherent-potential approximation[2,3] (CPA), now ending its first decade. Although efforts to improve our understanding beyond the CPA have not all met with the same good fortune, there have been recent exceptions. Cluster methods[4,5] have been devised which are accurate enough to reproduce the "peaky" structure of the density of states $\rho(\omega)$, which they sometimes do (notably in one dimension[4]) with startling fidelity. We have been working along such a cluster-type approach, and have found an extremely simple method translating directly into a computer algorithm. While unsuited to the theoretical study of Lifshitz[6] tails, our method has permitted us to reproduce many of the other known results over the theoretically permitted range of energy,[7] even near the energy maxima and minima, and additionally, permitted to study of the effect of short-range order. Along with Lifshitz, we envisage tails in $\rho(\omega)$ at the energy maxima and minima as arising from accidental correlations in increasingly large clusters, of a size that for practical reasons we are not at present capable of handling; however, the simplicity of the present method may suggest a natural extension to cover this.[8]

The basic outline of our paper is as follows: In Sec. II we present a method for the calculation of the single-body Green's function in the presence of an arbitrary number of impurities. We then discuss how our procedure can be implemented by the use of a convergence factor Σ. Section III is devoted to an analysis of the meaning and uses of the complex self-energy Σ within the context of a disordered medium. Results from our method are presented in Sec. IV, including the effects of short-range order. Beyond this in Sec. V we make further approximations that allow us to determine $\overline{G}_{\vec{k}\vec{k}}(\omega)$. Section VI is concerned with the development of a transport theory compatible with $\overline{G}$, along the lines of Baym and Kadanoff.[9]

II. CLUSTER GREEN'S FUNCTION

Let the Hamiltonian for the electrons within a single tight-binding band in a hypercubic lattice in D dimensions be

$$H=\sum_{i,j} T_{ij}\,|i\rangle\langle j| + \sum_i V_i\,|i\rangle\langle i| \equiv T+V\,, \qquad (1)$$

with $T_{ij}=(2D)^{-1}$ for i,j nearest neighbors and zero otherwise, $|i\rangle$ the Wannier state at the lattice point R_i, and V_i the potential which takes on one of two values depending whether atom A or B occupies the i site. We construct the resolvent operator $G(z)$ and its various matrix elements:

$$G(z)\equiv(z-H)^{-1}=[z-(T+\Sigma)-(V-\Sigma)]^{-1}\,, \qquad (2)$$

in which we reference the operators to a complex "optical potential" $\Sigma(z)$ merely as a device to enhance the convergence of subsequent expansions, with z equal to the frequency ω, extended to the complex plane.

For those readers familiar with the CPA, it is important to note that our new departure consists principally in dissociating the complex self-energy parameter $\Sigma(\omega)$ from the site-diagonal averaged Green's function $\overline{G}_{nn}(\omega)$. Whereas in CPA, knowledge of the one implies the other, via the relationship

$$\overline{G}_{nn_{\mathrm{CPA}}}(z)=\langle n|\{z-[T+\Sigma_{\mathrm{CPA}}(z)]\}^{-1}|n\rangle\,, \qquad (3)$$

our experience indicates that it is better to treat $\Sigma(\omega)$ merely as a convergence parameter, one to be chosen as an *ad hoc* aid in the calculations rather than by tedious and unnecessary self-consistency conditions. As by Eq. (2) the *exact* $\overline{G}_{nm}(z)$ are all independent of $\Sigma(z)$, in any *accurate approximation* to $\overline{G}_{nm}(z)$ we have latitude in our choice of $\Sigma(z)$, as discussed below, and we pick the simplest possible $\Sigma(z)$ for which our calculated G is approximately stationary.

We next define a modified resolvent operator $G^{(i)}$ appropriate to the case in which one sets $\tilde{V}_i = 0$, where we define $\tilde{V}_i \equiv \langle i | (V - \Sigma)_i | i \rangle$ and, indicating the elimination of the localized fluctuation potential at this site by $(\ )'_i$, we have

$$G^{(i)}(z) \equiv [z - (T + \Sigma) - (V - \Sigma)'_i]^{-1} . \tag{4}$$

The full resolvent (2) can be expressed in terms of the modification in (4) by the use of the operator identity $(A - B)^{-1} = A^{-1} + A^{-1}B(A - B)^{-1}$:

$$G(z) = G^{(i)}(z) + G^{(i)}(z)(V - \Sigma)_i G(z) . \tag{5}$$

Because the perturbation is diagonal in the Wannier representation, the matrix elements are easily found:

$$G_{nm}(z) = G_{nm}^{(i)}(z) + G_{ni}^{(i)}(z)\tilde{V}_i G_{im}^{(i)}(z) / (1 - G_{ii}^{(i)}(z)\tilde{V}_i) . \tag{6}$$

For the calculation of the density-of-states function $\rho(\omega) \equiv (-1/\pi)\,\mathrm{Im}\,\overline{G}_{nn}(\omega + i\epsilon)$ only the configurationally averaged Green's function $\overline{G}_{nn}(\omega + i\epsilon)$ is required. For the one-particle propagators $\overline{G}_{\vec{k}\vec{k}}$ the averaged Fourier transforms of all $\overline{G}_{nm}(\omega + i\epsilon)$ are needed. Equation (6) is now iterated. Define $\overline{G}^{(i,j)}(z)$ to be the modified resolvent operators with the fluctuation potentials $\tilde{V}$ at sites i and j removed. By a repetition of the above, we have

$$G_{nm}^{(i)}(z) = G_{nm}^{(i,j)}(z) + G_{nj}^{(i,j)}\tilde{V}_j G_{jm}^{(i,j)} / (1 - G_{jj}^{(i,j)}\tilde{V}_j) . \tag{7}$$

The matrix elements G_{nm} decay exponentially with distance R_{nm}; thus the expansions (6) and (7) are in a symbolic "parameter" γ defined as $G_{nj}^{(\cdots)}\tilde{V}_j$, which is "small" for small $\tilde{V}_j$ and "exponentially small" at large $\tilde{V}_j$. *The processes (6) and (7) are to be repeated any number of times, until the largest practical cluster size is achieved.*[10] Termination, by truncation, of the series consists of approximating the most distant G's, i.e., those with the largest number of superscripts, by their value in the average optical potential. Thus, if we stop at (7), the approximation consists in replacing $G_{nm}^{(i,j)}(z)$ by $\langle n | [z - (T + \Sigma)]^{-1} | m \rangle$. The configurational averages over all the explicitly retained V_i are then performed, and all $\overline{G}$'s obtained.

III. CHOICE OF Σ

We now come to our principal point of departure from other methods—our choice of Σ. Our results would depend crucially upon Σ except for the following observations. Since the behavior of the local cluster is the dominant characteristic of disordered systems, we expect results insensitive to the particular choice of Σ if the cluster size is sufficiently large.

We require a simple functional form for Σ that allows for states out to the bands limits. This excludes the use of Σ_{CPA} which is known to produce bands that are always too narrow. We restrict the range of possible Σ's by requiring that it obey dispersion relations, insuring that our approximate G is analytic. Furthermore, a functional form is desired in which G is accurate in both the weak as well as strong scattering regimes. Because of the local nature of highly disordered systems, our choice becomes more critical for small potential differences where effects are more extended. Our input is the $\mathrm{Im}\,\Sigma$ which we take as one or more step functions, nonzero only within the theoretical band limits. $\mathrm{Re}\,\Sigma$ is then determined from the following dispersion relation:

$$\Sigma(z) = \tfrac{1}{2}\langle V \rangle + \frac{1}{2\pi i}\int_{-\infty}^{\infty} dx\, \frac{\Sigma(x + i\epsilon)}{x - z} . \tag{8}$$

This is sufficient to make our approximation to $G(z)$ satisfy causality. The density-of-states sum rule, $\int_{-\infty}^{\infty} d\omega\, \rho(\omega) = 1$ is itself a beneficial consequence of the analyticity of our approximate $G(z)$ and its resulting $1/z$ dependence in the asymptotic limit as we discuss elsewhere.[11] We verified that in all cases studied, the sum rule on $\rho(\omega)$ was satisfied numerically. If we choose the constant $\mathrm{Im}\,\Sigma$ to be of magnitude of $\mathrm{Im}\,\Sigma_{\mathrm{CPA}}$, then in the weak-scattering and low-concentration regimes $G(z)$ will be quite similar to $G_{\mathrm{CPA}}(z)$ so that accurate results can be expected in all regimes.

Before we proceed, the way in which we use $\mathrm{Im}\,\Sigma_{\mathrm{CPA}}$ must be more clearly outlined. In the accompanying Fig. 1 we display the two basic behavior patterns of $|\mathrm{Im}\,\Sigma_{\mathrm{CPA}}|$ as observed by Velický, Kirkpatrick, and Ehrenreich.[12] It should be noted that, here also, $\mathrm{Re}\,\Sigma$ and $\mathrm{Im}\,\Sigma$ are related by the Eq. (8). Σ has to describe everything in the CPA; it determines band gaps, peaks in the density of states, and the general overall scale. Most of these results (e.g., band gaps and complicated structure) are better obtained by our detailed calculations of the correlated scattering. We hypothesize that the most useful information from CPA is contained in the general overall magnitude of $\mathrm{Im}\,\Sigma_{\mathrm{CPA}}$. Operationally, in Fig. 1(a), we would ignore values of $|\mathrm{Im}\,\Sigma_{\mathrm{CPA}}|$ from the region of its maximum as well as the extremeties of the band. In the former range of energies, we expect exceptional scattering because it is easiest for these states to make transitions due to band overlap, whereas at the band edges the spectrum will be least disturbed, according to the same considerations. Any value from the shaded region is then acceptable. In terms of particle lifetimes, we will obtain the large and small transition rates be-

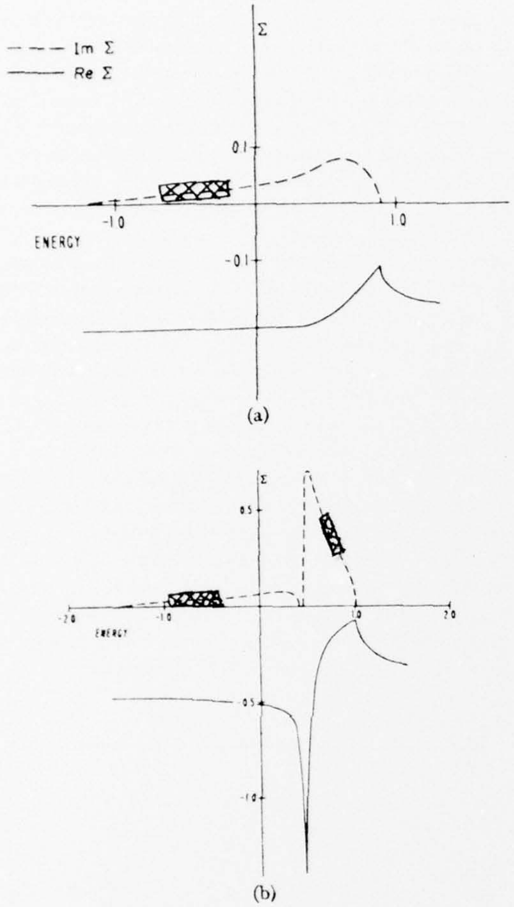

FIG. 1. Real and imaginary parts of the complex self-energy in the CPA. The two sets of curves are indicative of the type of results that can be expected from this approximation. (a) A situation where the alloy bands overlap; (b) the case of split bands. Energy is in units of half-bandwidth. This is an adaption of a figure from Ref. 12.

cause we almost solve the eigenvalue problem exactly for each configuration and this is clearly equivalent to a perturbation approach. As for concentration dependence in Σ_{CPA}, we will obtain correct behavior simply because we weigh each configuration by its appropriate probability. Thus we are able to include both the dynamical and statistical aspects of the problem.

In Fig. 1(b), the same analysis leads us to ignore the very large values of $|\mathrm{Im}\Sigma_{CPA}|$ in both subbands. Here though, the magnitudes are considerably different leading us to suspect that two different constants are needed. Further details of this case will be elucidated in the following examples.

IV. ANALYSIS OF RESULTS

We first consider the canonical one-dimensional tight-binding binary alloy for three different scattering strengths at a 50-50 concentration. Figure 2 compares the results of one-, three-, and five-cluster calculations for $\rho(\omega)$ when $V_i = \pm 0.5$ with exact results. We see in this example the development of the peaky structure associated with special clusters of atoms as our cluster size increases. Proceeding to a larger scattering strength ($V_i = \pm 1.0$), we expect that the local configurations will play a more prominent role because of increased wave function localization. As shown in Fig. 3 we successfully reproduce most of the structural details of $\rho(\omega)$ for a five cluster. To check the degree of insensitivity in our five-cluster model we varied $|\mathrm{Im}\Sigma|$ within the limits given by $\mathrm{Im}\Sigma_{CPA}$ and found little change in the overall pattern as shown in Fig. 4. This indicates, numerically that the resulting G is stationary and that Σ is optimum.

The scattering strengths are now increased to $V_i = \pm 2.0$, providing a critical evaluation of the methods capabilities (larger scattering strengths

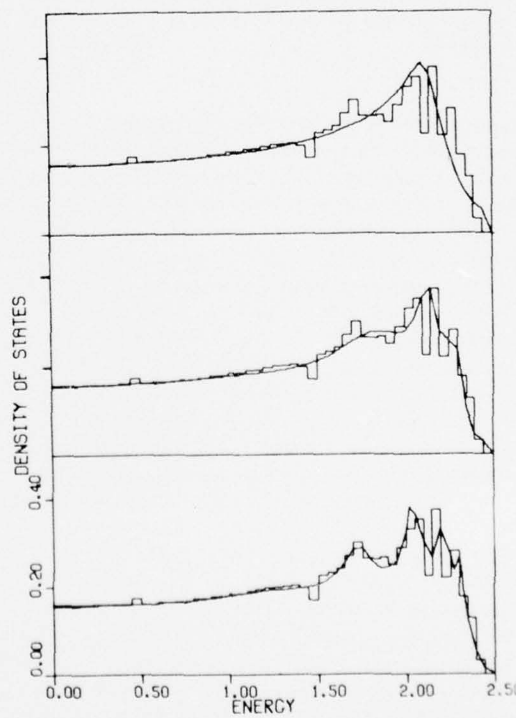

FIG. 2. Comparison of one-, three-, and five-cluster calculations for $\rho(\omega)$ using $|\mathrm{Im}\Sigma_{trial}| = 0.15$, with $V_i = \pm 0.5$, $c = 0.5$. Background (histogram) is exact results of Ref. 4. Energy is in units of half-bandwidth.

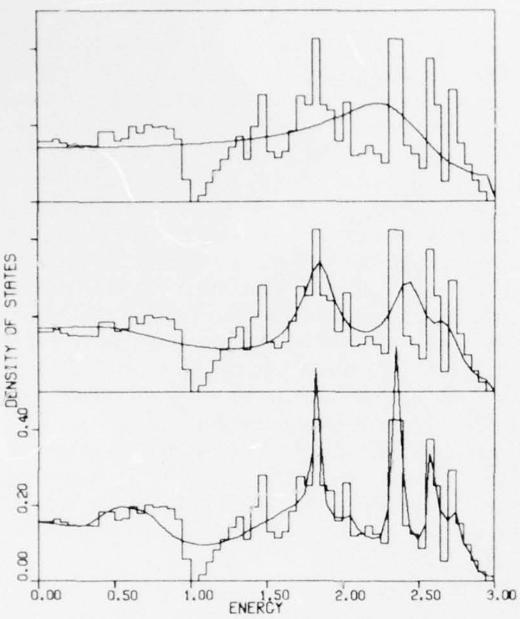

FIG. 3. Comparison of one-, three-, and five-cluster calculations for $\rho(\omega)$ in one dimension using $|\mathrm{Im}\Sigma_{\mathrm{trial}}| = 0.5$, with $V_i = \pm 1.0$ and $c = 0.5$. Background (histogram) is exact results of Ref. 4. Energy is in units of half-bandwidth.

are in a sense too easy because wave function localization makes a cluster calculation more plausible). Using the exact scattering off all configurations of five atoms, the highly discrete spectrum is well reproduced, as seen in Fig. 5(a). Increasing the cluster size to seven atoms, keeping the convergence factor the same as in Fig. 5(a), improves our agreement with the exact results as shown in Fig. 5(b). In Fig. 6 we display the results of again varying $|\mathrm{Im}\Sigma|$ within the limits dictated by $\mathrm{Im}\Sigma_{\mathrm{CPA}}$; the major details are again seen to remain stationary. We have found empirically that if $|\mathrm{Im}\Sigma|$ is too small, the resultant density of states is too "peaky" and as such, representative of a molecular cluster, instead of the solid state. If $|\mathrm{Im}\Sigma|$ is too large, then the central site predominates, as is correct only in the extreme "atomic" limit when potential fluctuations greatly exceed the bandwidth. One can see this from Fig. 6 since the sharper curve is associated with the lowest value of $|\mathrm{Im}\Sigma|$, and vice versa.

In three dimensions the obvious cluster size is seven sites. Figure 7 compares our calculation with the Monte-Carlo-type numerical results of Alben *et al.*[13] A constant $\mathrm{Im}\Sigma$ gave poor results in this case, but the CPA calculations immediately showed us why: $\mathrm{Im}\Sigma_{\mathrm{CPA}}$ was more than one order of magnitude smaller in the majority subband than in the minority subband. Consequently we changed $\mathrm{Im}\Sigma$ to the step function shown in the figure, varying the parameters (magnitudes of the steps) again guided by CPA. The results now agreed well with the exact computations and were insensitive to the precise value of our parameters as is evidenced by Fig. 8 in which a three-step function was used.

To illustrate entirely new applications, consider effects of short-range order on this same alloy. With α the Cowley short-range order parameter—c_A and $c_B = 1 - c_A$ the relative concentrations and P_{AB} the probability of finding atom A at a given site when a B atom occupies a specified neighboring site—we have $P_{AA} = c_A + c_B\alpha$, $P_{BA} = c_B(1-\alpha)$, $P_{AB} = c_A(1-\alpha)$, and $P_{BB} = c_B + c_A\alpha$. In Ref. 13, $\alpha = 0$. For $c_A = 0.1$, α can vary from -0.11 to $+1.0$; negative α is associated with enhanced tendency of A atoms to be surrounded by B's (i.e., "antiferromagnetism"), positive α indicates enhancement in the probability of either species being surrounded by atoms of its own kind (i.e., "ferromagnetism"). Using the same convergence parameters as in our calculation at $\alpha = 0$, in Fig. 9 we find distinctive features in the minority subband

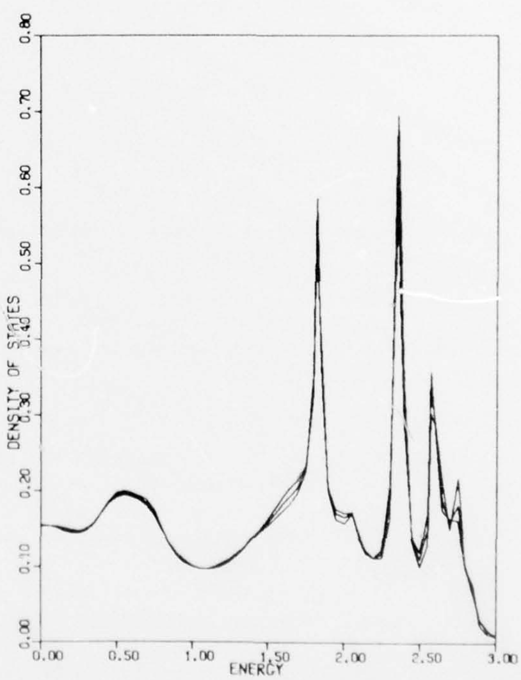

FIG. 4. Curves give $\rho(\omega)$ for three different values of Σ_{trial} for a five-cluster calculation of a one-dimensional alloy with $V_i = \pm 1.0$ and $c = 0.5$. The sharpest peaks are associated with lowest value of $|\mathrm{Im}\Sigma_{\mathrm{trial}}|$, 0.4. Other values are 0.5 and 0.6. Energy is in units of half-bandwidth.

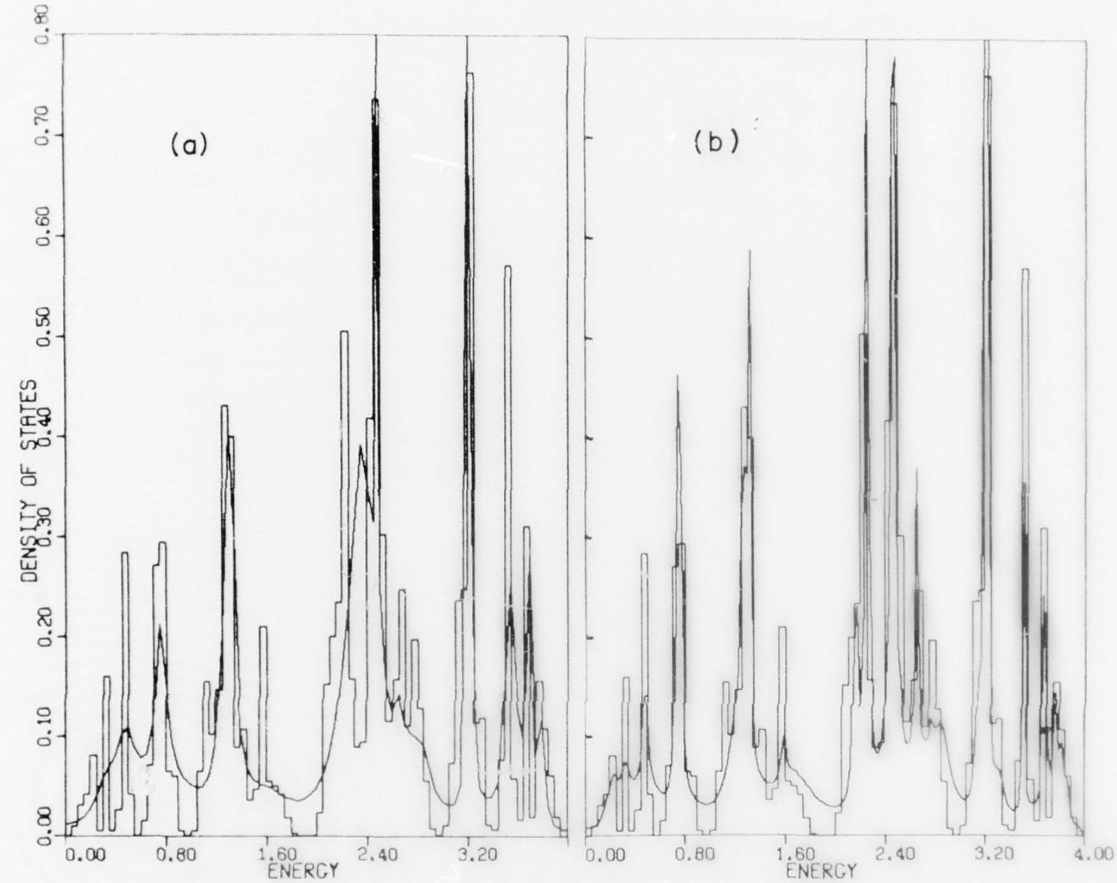

FIG. 5. (a) Density of states for a (50–50)% concentration one-dimensional alloy with $V_A = 2.0$, $V_B = -2.0$. Histogram is exact calculations from Ref. 4. The full band is obtained by reflecting the portion shown through the origin. These results were computed from $\mathrm{Im}\Sigma = -0.80$ and $\mathrm{Re}\Sigma$ obtained from $\mathrm{Im}\Sigma$ by Eq. (8), including the exact scattering from all configurations of a central atom and its four nearest neighbors. (b) Extension of the above results to a cluster of seven atoms using the same $\Sigma(\omega)$. This result is comparable in accuracy and wealth of detail to the best self-consistent calculation to date, Ref. 4. Energy is in units of half-bandwidth.

density of states that we interpret in terms of minority-atom clustering: the single peak of $\alpha = -0.07$ registers the unlikelihood of finding two A atoms as nearest neighbors, and the double peaks of $\alpha = 0.7$ represent the tendency of the same atoms to form pairs, triplets, etc. However, due to the sparseness of A atoms, triplets and higher-order clusters are statistically insignificant for these values of α.

V. ELECTRON PROPAGATION

So far we have developed a method for calculating the site-diagonal configuration averaged Green's function. We have not indicated, how we would calculate the non-site-diagonal propagators. One alternative is to develop a cluster method for the latter, similar to the method we used for the former. Another, simpler though less accurate, alternative will be employed. We first define a new self-energy $\Lambda^*(z)$ by the equation

$$\overline{G}_{ll}(z) = \frac{1}{N}\sum_{\vec{k}} \frac{1}{z - \Lambda^*(z) - \epsilon_{\vec{k}}}. \tag{9}$$

This relationship is numerically inverted to obtain Λ^* as a function of the exact or numerically calculated $\overline{G}_{ll}$. From the calculation of $\Lambda^*(\omega)$ we obtain $\Lambda^*(z)$ in the entire complex plane as

$$\Lambda^*(z) = \langle V\rangle + \frac{1}{\pi}\int_{-\infty}^{\infty} d\omega \frac{\mathrm{Im}\,\Lambda^*(\omega + i\epsilon)}{\omega - z}. \tag{10}$$

If further values of $\overline{G}_{lm}(z)$ were calculated numer-

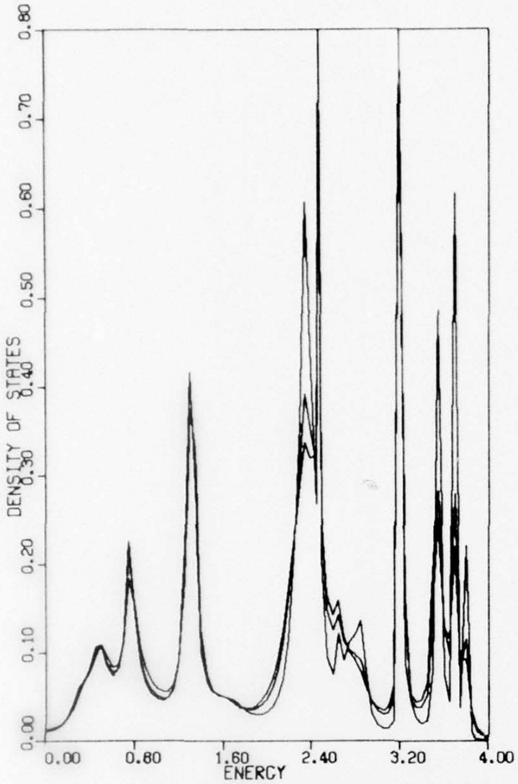

FIG. 6. Curves give $\rho(\omega)$ for three different values of Σ_{trial} for a five-cluster calculation of a one-dimensional alloy with $V_i = \pm 2.0$ and $c = 0.5$. The sharpest peaks are associated with lowest value of $|\text{Im}\,\Sigma_{\text{trial}}| = 0.4$, other values are 0.8 and 1.0. Energy is in units of half-bandwidth.

ically, then we would determine $\Lambda^*(\vec{k}, z)$ from

$$\bar{G}_{lm}(z) \equiv \frac{1}{N}\sum_{\vec{k}} \frac{e^{i\vec{k}\cdot(\vec{R}_l - \vec{R}_m)}}{z - \Lambda^*(\vec{k}, z) - \epsilon_{\vec{k}}} \tag{11}$$

and

$$\bar{G}_{\vec{k}\vec{k}} = [z - \Lambda^*(\vec{k}, z) - \epsilon_{\vec{k}}]^{-1} .$$

The analysis is facilitated by going over to a localized representation in which we would specify the number of elements $\Lambda^*_{lm}(z)$ that we have determined numerically. For example, if we have available $\bar{G}_{l,l}(z)$, $\bar{G}_{l,l+1}(z)$, and $\bar{G}_{l,l+2}(z)$, then we would be able to obtain

$$\Lambda^*_{l,m}(z) = \delta_{l,m}\Lambda^*_0(z) + \delta_{l+1,m}\Lambda^*_1(z) + \delta_{l+2,m}\Lambda^*_2(z) ,$$

by solving the three equations simultaneously. In the case at hand, we will use a site-diagonal self-energy since all we have at our disposal are the computed $\bar{G}_{ll}(z)$. We will still use

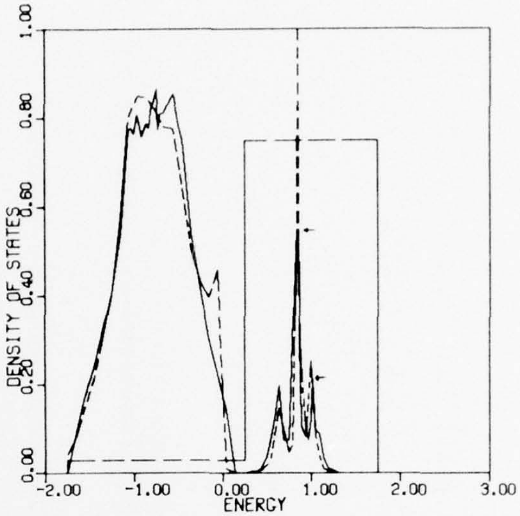

FIG. 7. Comparison of seven-cluster calculations of $\rho(\omega)$ (dashed line) using a two-step $|\text{Im}\,\Sigma_{\text{trial}}|$ (long dashed line), with numerical work of Alben *et al.* (Ref. 13) (solid line) who solved the Schrödinger equation for an 8000-atom three-dimensional tight-binding solid. The potentials are $V_i = \pm 0.75$ with impurity concentration of 0.1. Small horizontal arrows indicate the height to which their peaks rise. Energy is in units of half-bandwidth.

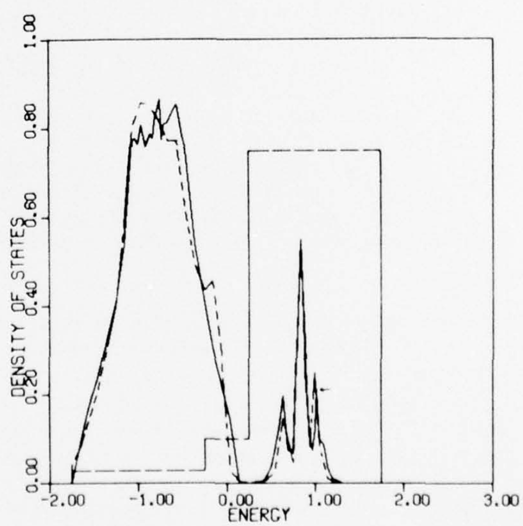

FIG. 8. Comparison of seven-cluster calculation of $\rho(\omega)$ (dashed line) using a three-step $|\text{Im}\,\Sigma_{\text{trial}}|$ (long dashed line) with results of Alben *et al.* (Ref. 13) (solid line). $V_i = \pm 0.75$ and $c = 0.1$ for this three-dimensional tight-binding alloy. The sharp peak in Fig. 7 at $\omega = 0.86$ is absent because of the coarser energy scale used. Arrow indicates the height to which their peak rises. Energy is in units of half-bandwidth.

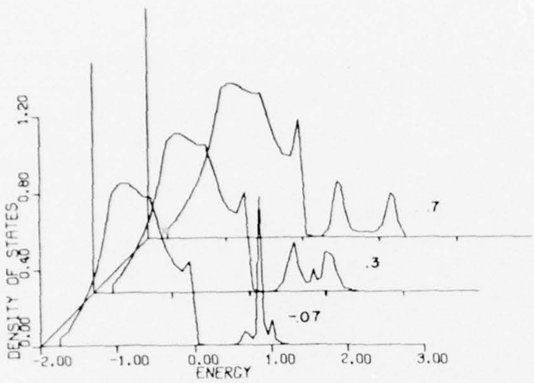

FIG. 9. Density of states for $V_A=0.75$, $V_B=0.75$ in a three-dimensional simple cubic lattice with concentration $c_A=0.1$. Cowley short-range order parameter α is -0.07, 0.3, and 0.7, respectively. $|\mathrm{Im}\Sigma|$ is the same as in Fig. 7 (dot-dash line) and $\mathrm{Re}\Sigma$ is obtained therefrom by use of Eq. (8). Energy is in units of half-bandwidth.

$$\overline{G}_{lm}(z)=\frac{1}{N}\sum_{\vec{k}}\frac{e^{i\vec{k}\cdot(\vec{R}_l-\vec{R}_m)}}{z-\Lambda^*(z)-\epsilon_{\vec{k}}},$$

hence (12)

$$\overline{G}_{\vec{k}\vec{k}}=[z-\Lambda^*(z)-\epsilon_{\vec{k}}]^{-1},$$

as the definition of the off-diagonal elements. The propagators decay rapidly with distance R_{lm} so both the one-, two-, or three-point curve-fitting procedures will probably give reasonably equivalent results. Now, all the information contained in our previous numerical work is stored in $\Lambda^*(z)$, the complex proper self-energy part. It is of interest to compare $\mathrm{Im}\Lambda^*$ with $\mathrm{Im}\Sigma_{CPA}$ in order to see how they differ. This is done in Fig. 10 for the one-dimensional alloy of Figs. 5 and 6.

In summary we have presented a relatively simple method for calculating the eigenvalue spectrum of a disordered system, one that avoids all the computational pitfalls of self-consistent methods. This quasi-invariant theory is not only highly accurate, but also allows the bounds on the frequency spectrum to be naturally determined by the correlated scattering of a local group. We now discuss transport and develop a formalism that allows our numerical output to be used in approximations that conserve particle number and energy.

VI. TRANSPORT IN DISORDERED SYSTEMS

The linear response of the current to the electric field defines the conductivity, which we take to be the same along the three principal directions in our simple cubic structure. Following Velický[14] we have in our single-particle model

$$\sigma(0)=\frac{\pi}{m^2}\int_{-\infty}^{\infty}d\lambda\left(-\frac{\partial f(\lambda)}{\partial\lambda}\right)\langle\!\langle\delta(\lambda-H)p_1\delta(\lambda-H)p_1\rangle\!\rangle, \tag{13}$$

where e, the electric charge, is unity and f is the fermi function. The bracketed term is short hand for

$$\langle\!\langle\delta(\lambda-H)p_1\delta(\lambda-H)p_1\rangle\!\rangle \equiv\sum_\alpha\langle\!\langle\alpha|\delta(\lambda-H)p_1\delta(\lambda-H)p_1|\alpha\rangle\!\rangle. \tag{14}$$

p_1 is the momentum operator along an arbitrarily chosen principal axis and $\langle\ \rangle$ denotes configuration averaging. Examination of Eq. (14) reveals that we require the two-particle correlation function

$$\langle G^2_{imjl}(z_1,z_2)\rangle=\langle\!\langle i|(z_1-H)^{-1}|j\rangle\langle m|(z_2-H)^{-1}|l\rangle\!\rangle. \tag{15}$$

We can relate $\langle G^2\rangle$ to $\overline{G}$ by the equation

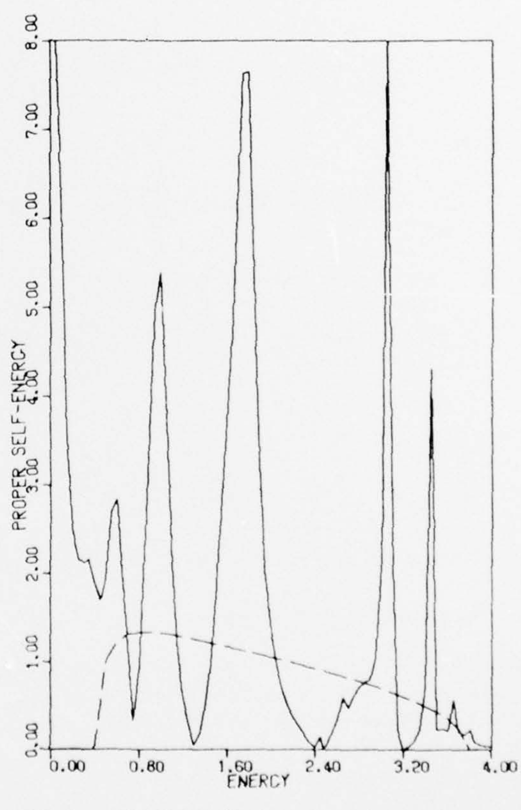

FIG. 10. Solid line is $|\mathrm{Im}\Lambda^*(\omega)|$ for the one-dimensional alloy of Figs. 5 and 6 in the five-cluster approximation while the dashed line is the corresponding $|\mathrm{Im}\Sigma_{CPA}(\omega)|$. Energy is in units of half-bandwidth.

$$\langle G^2_{imjl}(z_1, z_2)\rangle = \overline{G}_{ij}(z_1)\overline{G}_{ml}(z_2) + \sum_{\substack{kp\\qr}} \overline{G}_{ik}(z_1)\,\overline{G}_{ql}(z_2)\Xi_{kr,pq}\langle G^2_{pmjr}(z_1, z_2)\rangle\,, \tag{16}$$

which defines the index structure of the vertex function. This equation is exact for the exact $\overline{G}$ and we will use it to define $\langle G^2\rangle$ when we have an approximate single-particle Green's function. We can place restrictions on possible vertex functions by requiring the conservation of charge and energy in the presence of a long-wavelength disturbance. This leads to the introduction of a new operator

$$K(z_1, z_2) = \langle (z_1 - H)^{-1}(z_2 - H)^{-1}\rangle\,. \tag{17}$$

It is easy to show that K must satisfy the following Ward-type identity:

$$K(z_1, z_2) = (z_2 - z_1)^{-1}[\overline{G}(z_1) - \overline{G}(z_2)]\,, \tag{18}$$

which implies a connection between the one-body operator $\overline{G}$ and the two-body Green's function $\langle G^2\rangle$. Also, one can show with the above condition that the linear response of the particle number and energy to a long-wavelength disturbance is zero, thus ensuring the appropriate conservation laws. If we use an approximate $\overline{G}$, then we must construct a K that maintains Eq. (18) and this allows us to relate $\overline{G}$ to the vertex function in the following way: We let $z_2 \to z_1$, then Eq. (18) becomes

$$K_{il}(z_1, z_1) = -\frac{d\overline{G}_{il}(z_1)}{dz_1}\,. \tag{19}$$

The configuration averaged resolvent can be written

$$\overline{G}_{il}(z_1) = \langle i\,|\,[z_1 - T - \Lambda^*_{op}(z_1)]^{-1}\,|\,l\rangle\,, \tag{20}$$

which leads to the equation

$$K_{il}(z_1, z_1) = \sum_j \overline{G}_{ij}(z_1)\overline{G}_{jl}(z_1) + \sum_{kq} \overline{G}_{ik}(z_1)\left(\frac{-d\Lambda^*_{kq}(z_1)}{dz_1}\right)\overline{G}_{ql}(z_1)\,. \tag{21}$$

But

$$\frac{-d\Lambda^*_{kq}(z_1)}{dz_1} = \frac{1}{N}\sum_{rp}\frac{d\Lambda^*_{kq}(z_1)}{d\overline{G}_{pr}(z_1)}\left(\frac{-d\overline{G}_{pr}(z_1)}{dz_1}\right) = \frac{1}{N}\sum_{rp}\frac{d\Lambda^*_{kq}(z_1)}{d\overline{G}_{pr}(z_1)}K_{pr}(z_1, z_1)\,,$$

so that

$$K_{il}(z_1, z_1) = \sum_j \overline{G}_{ij}(z_1)\overline{G}_{jl}(z_1) + \sum_{\substack{kq\\rp}} \overline{G}_{ik}(z_1)\overline{G}_{ql}(z_1)\,\frac{1}{N}\,\frac{d\Lambda^*_{kq}(z_1)}{d\overline{G}_{pr}(z_1)}K_{pr}(z_1, z_1)\,. \tag{22}$$

In $\langle G^2\rangle$, we let $m = j$ and sum over all j with $z_1 = z_2$:

$$\sum_j \langle G^2_{ijjl}(z_1, z_1)\rangle = \sum_j \overline{G}_{ij}(z_1)\overline{G}_{jl}(z_1) + \sum_{\substack{kq\\rp}} \overline{G}_{ik}(z_1)\overline{G}_{ql}(z_1)\Xi_{kr,pq}(z_1, z_1)\sum_j \langle G^2_{pjjr}(z_1, z_1)\rangle\,. \tag{23}$$

Once we recognize

$$K_{il}(z_1, z_1) = \sum_j \langle G^2_{ijjl}(z_1, z_1)\rangle\,,$$

we find that the vertex must satisfy the equation

$$\Xi_{kr,pq}(z_1, z_1) = \frac{1}{N}\,\frac{d\Lambda^*_{kq}(z_1)}{d\overline{G}_{pr}(z_1)}\,. \tag{24}$$

We note that not only must this relationship hold for the exact vertex function and self-energy but also in any approximation in which it is desired that the two-particle correlation function satisfy the Ward identity, Eq. (18). This means that once an approximation is made to Λ^*, we can determine transport functions that allow the conservation laws to be obeyed. We could make further approximations to $\langle G^2\rangle$ but we would then have no guarantee of conserving charge, energy, etc. The Ward identity is useful to generate a unique compatible vertex function only when the frequencies are the same. For the case at hand, $\Lambda^*_{kq}(z_1) = \delta_{kq}\Lambda^*(z_1)$ is a function of the site-diagonal averaged Green's function so the vertex is

$$\Xi_{kr,pq}(z_1, z_1) = \frac{\delta_{kq}\delta_{rp}}{N}\,\frac{d\Lambda^*(z_1)}{dG_{rr}(z_1)} = \delta_{kq}\delta_{rp}\Xi(z_1, z_1)\,. \tag{25}$$

For $z_1 \neq z_2$, we make the approximation

$$\Xi_{kr,pq}(z_1, z_2) = \delta_{kq}\delta_{rp}\Xi(z_1, z_2)\,. \tag{26}$$

This is certainly consistent with the Ward identity, and furthermore, it allows us to show that contributions to the conductivity from the vertex corrections then vanish. The two-particle correlation function is now

$$\langle G^2_{imjl}(z_1,z_2)\rangle = \bar G_{ij}(z_1)\bar G_{ml}(z_2) + \sum_k \bar G_{ik}(z_1)\bar G_{kl}(z_2)\Xi(z_1,z_2)\sum_r \langle G^2_{rmjr}(z_1,z_2)\rangle . \tag{27}$$

To find the conductivity we need $\langle\langle \delta(\lambda_1-H)p_1\delta(\lambda_2-H)p_2\rangle\rangle$ or

$$I_{12}(\lambda_1,\lambda_2)=\sum_{ijlm} p^{\vec 1}_{jm}\langle\langle m|\delta(\lambda_2-H)|l\rangle\rangle p^{\vec 2}_{li}\langle\langle i|\delta(\lambda_1-H)|j\rangle\rangle , \tag{28}$$

which requires

$$\sum_{ijlm} p^{\vec1}_{jm}p^{\vec2}_{li}\langle G^2_{imjl}(\lambda_1,\lambda_2)\rangle = \sum_{ijlm} p^{\vec1}_{jm}p^{\vec2}_{li}\bar G_{ij}(\lambda_1)\bar G_{ml}(\lambda_2) + \sum_{ijlm}\sum_k p^{\vec2}_{li}\bar G_{ik}(\lambda_1)\bar G_{kl}(\lambda_2)\Xi(\lambda_1,\lambda_2)\sum_r\langle G^2_{rmjr}(\lambda_1,\lambda_2)\rangle p^{\vec1}_{jm} . \tag{29}$$

The second term breaks up into

$$\sum_{ilk} p^{\vec2}_{li}\bar G_{ik}(\lambda_1)\bar G_{kl}(\lambda_2)\sum_{jmr}\Xi(\lambda_1,\lambda_2)\langle G^2_{rmjr}(\lambda_1,\lambda_2)\rangle p^{\vec1}_{jm} \equiv AB . \tag{30}$$

Let us transform the Wannier sum in A to a Bloch sum. Then since $\langle \vec k|p_2|\vec k'\rangle = mV_2(\vec k)\delta_{\vec k\vec k'}$, and $\bar G$ is diagonal in the Bloch representation,

$$A = m\sum_{\vec k} V_2(\vec k)\bar G_{\vec k\vec k}(\lambda_1)\bar G_{\vec k\vec k}(\lambda_2) = m\sum_{\vec k}\frac{\partial\epsilon_{\vec k}}{\partial k_2}\,\frac{1}{\lambda_1-\Lambda^*(\lambda_1^+)-\epsilon_{\vec k}}\,\frac{1}{\lambda_2-\Lambda^*(\lambda_2^+)-\epsilon_{\vec k}} . \tag{31}$$

The propagators are even under inversion ($\vec k \to -\vec k$) but the velocity $V_2 = \partial\epsilon/\partial k_2$ is odd, giving us zero, and all vertex corrections now vanish. In this case, the fortunate cancellation of vertex corrections comes about as a consequence of the approximation of the proper self-energy by a site-diagonal quantity, Eq. (12).

VII. ZERO-TEMPERATURE dc CONDUCTIVITY

We are now in a position to evaluate $\sigma(0)$. Because the vertex corrections vanish,

$$I_{12}(\lambda,\lambda)=\sum_{ijlm}p^{\vec1}_{jm}\langle\langle m|\delta(\lambda-H)|l\rangle\rangle p^{\vec2}_{li}\langle\langle i|\delta(\lambda-H)|j\rangle\rangle = m^2\sum_{\vec k}V_1(\vec k)V_2(\vec k)\langle\langle\vec k|\delta(\lambda-H)|\vec k\rangle\rangle^2 . \tag{32}$$

With the definition

$$\langle\langle\vec k|\delta(\lambda-H)|\vec k\rangle\rangle = (-1/\pi)\,\mathrm{Im}\bar G_{\vec k\vec k}(\lambda^+),$$

we get for $\vec 1 = \vec 2$ $[\sigma(\omega)=\sigma_{11}(\omega)]$,

$$\sigma(0)=\frac{2}{\pi}\int_{-\infty}^{\infty}d\lambda\left(\frac{-\partial f(\lambda)}{\partial\lambda}\right) \times \sum_{\vec k} V_1(\vec k)^2[\mathrm{Im}\bar G_{\vec k\vec k}(\lambda^+)]^2 , \tag{33}$$

where we have included a factor of 2 for the two possible spin orientations. At $T=0$, $-\partial f(\lambda)/\partial\lambda = \delta(\lambda-\mu)$, with μ the chemical potential, and the conductivity per atom is

$$\sigma(0)=\frac{2}{\pi N}\sum_{\vec k} V_1(\vec k)^2[\mathrm{Im}\bar G_{\vec k\vec k}(\mu^+)]^2 \tag{34}$$

or

$$\sigma(0)=\frac{2}{\pi}\int_{-\infty}^{\infty}dE\left(\frac{\mathrm{Im}\Lambda^*(\mu^+)}{[\mu-\mathrm{Re}\Lambda^*(\mu^+)-E]^2+\mathrm{Im}\Lambda^*(\mu^+)^2}\right)^2 \times\frac{1}{N}\sum_{\vec k} V_1(\vec k)^2\delta(E-\epsilon_{\vec k}) . \tag{35}$$

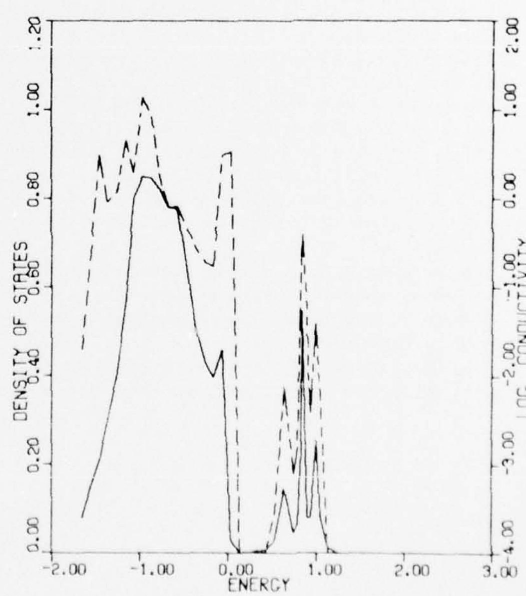

FIG. 11. Then dc conductivity (dashed line) of a three-dimensional alloy with $V_i = \pm 0.75$ and $c = 0.1$. Density of states taken from Fig. 7 is shown in the solid line. Energy is in units of half-bandwidth.

This natural separation, only possible for a proper self-energy independent of $\vec{k}$, isolates the lifetime and energy shifts of the single-particle excitations from that part of the conductivity which pertains to the particular lattice under study. We will concentrate on a three-dimensional simple cubic lattice with

$$\epsilon_{\vec{k}} = \tfrac{1}{3}(\cos k_x + \cos k_y + \cos k_z) .$$

Then,

$$V_1(\vec{k})^2 = \tfrac{1}{9}\sin^2 k_x = \tfrac{1}{9}(1 - \cos^2 k_x) .$$

Consider the functions

$$P_n(E) = \frac{1}{N}\sum_{\vec{k}} \frac{e^{nik_x}}{E - \epsilon_{\vec{k}}} ,$$

for which we have

$$\frac{1}{N}\sum_{\vec{k}} V_1(\vec{k})^2\delta(E - \epsilon_{\vec{k}}) = \frac{1}{18\pi}\,\mathrm{Im}[P_2(E^+) - P_0(E^+)] .$$

The imaginary part of both $P_2(E^+)$ and $P_0(E^+)$ vanish outside the unperturbed band and

$$\sigma(0) = \frac{1}{9\pi^2}\,\mathrm{Im}\int_{-1}^{1} dE \left(\frac{\mathrm{Im}\Lambda^*(\mu^+)}{[\mu - \mathrm{Re}\Lambda^*(\mu^+) - E]^2 + \mathrm{Im}\Lambda^*(\mu^+)^2}\right)^2 [P_2(E^+) - P_0(E^+)] . \qquad (36)$$

We have calculated the dc conductivity for our three-dimensional alloy in order to illustrate our formal results. Generally, there are two ways in which the dc conductivity can vanish. If the density of states at the Fermi level is zero then so is $\sigma(0)$. In addition, we can have a finite $\rho(\mu)$, but a zero mobility because of wave-function localization. Equation (36) only admits a zero in $\sigma(0)$ if $\rho(\mu)$ is zero so we cannot take the latter possibility into account. The conductivity is displayed in Fig. 11 against its respective density of states. There is a rather direct correlation between the magnitude of the density of states and that of the conductivity. This relationship is understood in terms of the availability of states at a given energy to which an initial state can make a transition. We also find peaks in $\sigma(0)$ which we associate with velocity peaks in the cubic band structure.

---

*Supported by AFOSR Grant No. 73-2430B.

†Research is supported by the ONR under Grant No. N00014-76-C-0690.

[1] R. J. Elliott, J. A. Krumhansl, and P. L. Leath, Rev. Mod. Phys. 46, 465 (1974); H. Ehrenreich and L. M. Schwartz, Solid State Phys. 31, 149 (1976); and J. D. Joannopoulos and Marvin L. Cohen, *ibid.* 31, 71 (1976).

[2] P. Soven, Phys. Rev. 156, 809 (1967); 178, 1136 (1969).

[3] D. W. Taylor, Phys. Rev. 156, 1017 (1967).

[4] W. H. Butler, Phys. Rev. B 8, 4499 (1973).

[5] Pabitra N. Sen and Felix Yndurain, Phys. Rev. B 13, 4387 (1976).

[6] R. Friedberg and J. M. Luttinger, Phys. Rev. B 12, 4460 (1975), and references cited therein.

[7] B. Velický, S. Kirkpatrick, and H. Ehrenreich, Phys. Rev. B 1, 3250 (1970).

[8] For example, while computer time increases exponentially with the number of neighboring shells over which one chooses to perform the exact configurational averages, the use of a Monte Carlo sampling technique may extend our calculations to large numbers of such shells at only a modest increase in computational time.

[9] G. Baym and L. Kadanoff, Phys. Rev. 124, 287 (1961).

[10] Our equations (6) and (7) were first derived by Vijay Kumar and S. J. Joshi [J. Phys. C 8, L148 (1975)] in the context of applying Butler's (Ref. 4) self-consistent equations to a tetrahedral alloy.

[11] A more complete exposition of this entire work is contained in the Ph.D. thesis of one of us (P.B.) (Yeshiva University, 1977) (unpublished).

[12] B. Velický, S. Kirkpatrick, and H. Ehrenreich, Phys. Rev. 175, 747 (1968).

[13] R. Alben, M. Blume, H. Krakauer, and L. Schwartz, Phys. Rev. B 12, 4090 (1975).

[14] B. Velický, Phys. Rev. 184, 614 (1969).